GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2000, 18:35:10 ; Search time 23795.2 seconds
                                                 (without alignments)
                                                 967.486 Million cell updates/sec

Title: US-09-227-881-34
Perfect score: 5271
Sequence: 1 atcttgttcagttacctc..........tcgggcatgagccagcaagg 5271

Scoring table: IDENTITY_NUC
                       Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
  1: gb_ba1:*
  2: gb_ba2:*
  3: gb_om:*
  4: gb_ov:*
  5: gb_pat:*
  6: gb_ph:*
  7: gb_pl1:*
  8: gb_pl2:*
  9: gb_pr1:*
 10: gb_pr2:*
 11: gb_pr3:*
 12: gb_ro:*
 13: gb_sy:*
 14: gb_un:*
 15: em_fun:*
 16: em_hum1:*
 17: em_hum2:*
 18: em_in:*
 19: em_om:*
 20: em_or:*
 21: em_ov:*
 22: em_pat:*
 23: em_ph:*
 24: em_pl:*
 25: em_ro:*
 26: em_sts:*
 27: em_sy:*
 28: em_un:*
 29: em_vi:*
 30: gb_ba3:*
 31: gb_in1:*
 32: gb_in2:*
 33: gb_in3:*
 34: gb_pl3:*
 35: gb_pr4:*
 36: em_ba1:*
 37: em_ba2:*
 38: em_htg1:*
 39: em_htg2:*
 40: em_htg3:*
 41: em_htg4:*
 42: em_htg5:*
 43: em_htg6:*
 44: em_htg7:*
 45: em_htg8:*
 46: em_htg9:*
 47: em_htg10:*
 48: em_htg11:*
 49: em_htg12:*
 50: em_htg13:*
 51: em_htg14:*
 52: em_htg15:*
 53: em_htg16:*
 54: em_htg17:*
 55: em_htg18:*
 56: em_htg19:*
 57: em_htg20:*
 58: em_htg21:*
 59: em_htg22:*
 60: em_htg23:*
 61: em_hum3:*
 62: em_hum4:*
 63: em_hum5:*
 64: em_hum6:*
 65: gb_pr5:*
 66: gb_pr6:*
 67: gb_pr7:*
 68: gb_htg1:*
 69: gb_htg2:*
 70: gb_htg3:*
 71: gb_htg4:*
 72: gb_htg5:*
 73: gb_htg6:*
 74: gb_htg7:*
 75: gb_htg8:*
 76: gb_htg9:*
 77: gb_htg10:*
 78: gb_htg11:*
 79: gb_htg12:*
 80: gb_htg13:*
 81: gb_htg14:*
 82: gb_htg15:*
 83: gb_htg16:*
 84: gb_htg17:*
 85: gb_htg18:*
 86: gb_htg19:*
 87: gb_htg20:*
 88: gb_htg21:*
 89: gb_htg22:*
 90: gb_htg23:*
 91: gb_sts1:*
 92: gb_sts2:*
 93: gb_vi1:*
 94: gb_vi2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		5271	100.0	5300	11	AF007562	AF007562 Homo sapi
2	c	5232.4	99.3	79376	65	HS454G6	298750 Human DNA s
3		5181.4	98.3	170425	77	AC024490	AC024490 Homo sapi
4		1804.4	34.2	2800	66	HSMYOC1	AF049791 Homo sapi
5		394.4	7.5	1086	66	HSGLC1A1	Z97171 Homo sapien
6		394.4	7.5	1228	9	AB006865S1	AB006686 Homo sapi
7	c	189.8	3.6	161577	10	AC007688	AC007688 Homo sapi
8	c	189.8	3.6	193123	77	AC023790	AC023790 Homo sapi
9	c	189	3.6	199722	71	AC012404	AC012404 Homo sapi
10		189	3.6	97037	67	AC004973	AC004973 Homo sapien
11	c	187.8	3.6	135038	67	HUMYWXD703	L78810 Homo sapi
12	c	185.8	3.5	76727	65	HS821D11	AL021453 Human DNA

```
RESULT    1

LOCUS       AF007562    5300 bp    DNA                      PRI    18-MAR-1998
DEFINITION  Homo sapiens trabecular meshwork inducible glucocorticoid response
            protein (TIGR) gene, promoter region and partial mRNA sequence.
ACCESSION   AF007562
VERSION     AF007562.1    GI:2970123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5300)
  AUTHORS   Nguyen,T.D., Chen,P., Huang,W.D., Chen,H., Johnson,D. and
            Polansky,J.R.
  TITLE     Gene structure and properties of TIGR, an olfactomedin-related
            glycoprotein cloned from glucocorticoid-induced trabecular meshwork
            cells
  JOURNAL   J. Biol. Chem. 273 (11), 6341-6350 (1998)
  MEDLINE   98165818
REFERENCE   2  (bases 1 to 5300)
  AUTHORS   Nguyen,T.D., Chen,P., Chen,H. and Polansky,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1997) Ophthalmology, University of California San
            Francisco, 10 Kirkham Street, San Francisco, CA 94143-0730, USA
FEATURES             Location/Qualifiers
     source          1..5300
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1..>5300
                     /gene="TIGR"
     promoter        1..5271
                     /gene="TIGR"
     mRNA            5272..>5300
                     /gene="TIGR"

                              ALIGNMENTS

13   185.8  3.5  152044  79  AC026395  Homo sapi    AC026395  Homo sapi
14   185.8  3.5  157057  78  AC025947  Homo sapi    AC025947  Homo sapi
15   185.8  3.5  161499  72  AC015488  Homo sapi    AC015488  Homo sapi
16   185.8  3.5  184656  86  AL161726  Homo sapi    AL161726  Homo sapi
17   185.8  3.5  200681  69  AC008755  Homo sapi    AC008755  Homo sapi
18   185.6  3.5  201372  86  AL157941  Homo sapi    AL157941  Homo sapi
19   185.8  3.5  146190  89  AP001813  Homo         AP001813  Homo sapi
20   184.8  3.5  163494  89  AP002391  Homo         AP002391  Homo sapi
21   184.8  3.5  183241  89  AP001354  Homo         AP001354  Homo sapi
22   184.6  3.5  176029  11  AC011362  Homo sapi    AC011362  Homo sapi
23   184.2  3.5  130020  67  AC004525  Homo sapi    AC004525  Homo sapi
24   184    3.5  157304  78  AC024720  Homo sapi    AC024720  Homo sapi
25   184    3.5  187709  73  AC016168  Homo sapi    AC016168  Homo sapi
26   184    3.5  233734  70  AC011407  Homo sapi    AC011407  Homo sapi
27   183.6  3.5   62070  88  AL358852  Homo sapi    AL358852  Homo sapi
28   183.6  3.5  129370  89  AP000609  Homo sapi    AP000609  Homo sapi
29   183.6  3.5  169333  89  AP002789  Homo sapi    AP002789  Homo sapi
30   183.6  3.5  193171  74  AC018723  Homo sapi    AC018723  Homo sapi
31   183.6  3.5   41407  11  AC053467  Homo sapi    AC053467  Homo sapi
32   183.6  3.5  149138  79  AC026936  Homo sapi    AC026936  Homo sapi
33   183.4  3.5  156331  90  HSJ193M11 Homo sap     HSJ193M11 Homo sap
34   183.4  3.5  174974  81  AL356055  Homo sap     AL356055  Homo sapi
35   183.4  3.5  195364  65  HS431A14  Homo sap     HS431A14  Homo sapi
36   183.4  3.5  198583  74  AC019114  Homo sapi    AC019114  Homo sapi
37   183.4  3.5   41936  35  CH19R30879 Homo        CH19R30879 Homo sap
38   183.2  3.5   70128  10  HS267M20  Homo sap     HS267M20  Homo sapi
39   183.2  3.5  119483   9  AC005588  Homo sap     AC005588  Homo sapi
40   183.2  3.5  145528   6  AC003665  Homo         AC003665  Homo sapi
41   183.2  3.5  167943  65  AC005588  Homo sap     AC005588  Homo sapi
42   183.2  3.5  170245  65  HS109F14  Homo         HS109F14  Homo sap
43   183    3.5   42686   9  AC000093  Homo sapi    AC000093  Homo sapi
44   183    3.5   91767  10  AC007748  Homo sap     AC007748  Homo sapi
45   183    3.5  112748  10  AC007242  Homo sap     AC007242  Homo sapi
```

```
/note="additional 3' sequence deposited as TIGR mRNA with
GenBank Accession Number U85257"
/product="trabecular meshwork inducible glucocorticoid
response protein"

BASE COUNT     1483 a    1152 c    1236 g    1429 t
ORIGIN

Query Match     100.0%;  Score 5271;  DB 11;  Length 5300;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 5271;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy         1 atctttgttcaagttcacctcaagggctattatgaaatgaaatgagataac caatac   60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATCTTTGTTCAAGTTCACCTCAAGGGCTATTATGAAATGAAATGAGATAAC CAATAC   60

Qy        61 tcctataaactgtatagcctcatccggatgtatgtcttcggaagtgaataagaatca  120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TCCTATAAACTGTATAGCCTCATCCGGATGTATGTCTTCGGAAGTGAATAAGAATCA  120

Qy       121 ggaagaaggagtgtattccacgttagccaagtgtccaggctgttgtctgtctttagtga  180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GGAAGAAGGAGTGTATTCCACGTTAGCCAAGTGTCCAGGCTGTTGTCTGTCTTTAGTGA  180

Qy       181 cagatgtgctcctgacagaagctattcttcctccaggaaaaccacatccaatatggtaaatc  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CAGATGTTGCTCCTGACAGAAGCTATTCTTCAGGAAACATCACAATCCAATATGGTAAATC  240

Qy       241 catcaaaacgaggagctaagacacaggaagaaatgagatgggcacttgccaagaaaaatgccag  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CATCAAAACAGGAGCTAAGAAACAGGAGATGGGCACTTGCCCAAGGAAAATGCCAG  300

Qy       301 gagagacaaataatgatgaaaaataaacattttccctttgtttttaattcaggaaaatg  360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GAGAGACAAATAATGATGAAAAAATAAACTTTTCCCTTTGTTTTTAATTTCAGGAAAATG  360

Qy       361 atgaggaccaaaatcaatgaataaggaaaacagcctcagaaaagatgtttccaaatgg  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ATGAGGACCAAAATCAATGAATAAGGAAAACAGCTCAGAAAAGATGTTTCCAAATGG  420

Qy       421 taattaagtattttcccttgggaagagacctccatgtgagccttgagtcggggaaatggggaa  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TAATTAAGTATTTGTTCCTTGGGAGGAGACCTCCATGTGAGCTTGAGTCGGGGAAATGGGGAA  480

Qy       481 aaacgtcaaaagcagcatgatctgatcagatcccaaagtggattattattttaaaaccagat  540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AAACGTCAAAAGACATGATCTGATCAGATCCCAAAGTGGATTATTATTTTAAAAACCAGAT  540

Qy       541 ggcatcacctcctgggaggcaagtcatgttagcaaaggaccaacaatac  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GGCATCACTCTGGGGAGGCAAGTCATGTTAGCAAAGGACCAACAATAC  600

Qy       601 agcaaaatcaaattccgaaatcaaggagaggaggaaagctttcatac  660
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       601 AGCAAAATCAAATTCCGGAAATCAAGGAGGAGGAAAGCTTTCATAC  660

Qy       661 agtgattaggcagttgaccagtgcctcaaagacagcagtgccctcagca  720
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AGTGATTAGGCAGTTGACCATGTTCGCAACACCTCCTATACCAGGAACACAAA  720

Qy       721 attgaactgggctaagcctgggaaatatgaaaactgagagcaaacacaaa  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 ATTGACTGGGCTAAGCCTGGGAACTTCAAGGAAATATGAGAGCAAACACAAA  780

Qy       781 gacatggttaaaggcaaccagacaatctgtgagccttcaaagcagtgccctcagca  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GACATGGTTAAAAGGCAACACCAGACAACATGTGACCCTGCCCCTCAGCA  840

Qy       841 gggaccctgagcgcattgcctttaggaagccagtctttaaggaaatcttaagaaactc  900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GGGACACCTGAGCCATTGCCTTTAGGAAGGCCAGTTTTGCCTTTAGGAAGGAATGTTTAAGGAAATTTCC  900

Qy       901 ttgaaagtcatgaatttaaccattaagtatataaacaaatatcggatcataatcag  960
```

```
Db   901  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  960
Qy   961  TTGAAAGATCANTGAATTTAACCATTTAAAGTATAAAAACAAATATGGATGGATAATCAG 1020
Db   961  TTTAGACATGGGTCCCAAATTTTATAAAGTCAGGCATCAAGGATAACGTGTCCAGCTCC   1020

Qy  1021  ggataggttcagaaatcattgaaaatcactgtgtccccatcctaactttttccagaatgatc 1080
Db  1021  GGATAGGTTCAGAAATCATTAGAAATCACTGTGTCCCCATCCTAACTTTTTCAGAATGATC 1080

Qy  1081  tgtcatagtgcctcacacacaggcccgatgtgtctgacctactaacaatctacaaaccaa  1140
Db  1081  TGTCATAGGCCCTCACACACAGGCCCGATGTGTCTGACCTACTAACAATCTACAACCCAA   1140

Qy  1141  gtgcctcaaagccaattgttcatctcagtaggtctgacctaactaacctacaacccccc   1200
Db  1141  GTGCCTCAACCATTGTTAACGTGTCATCTCAGTAGGTCCCAATTACAAAATGCCACCTCCC 1200

Qy  1201  tgtgcagcccatcccgcctcccacactctgcatcacgatgt                     1260
Db  1201  TGTGCAGCCCATCCCGCCTCCCATCCTAACTTTTTCAGAATGATC                 1260

Qy  1261  tacagccagaagctccgtgaggtgaggtctgttgtctcttacccctacctgtatgtctac 1320
Db  1261  TACAGCCAGAAGCTCCGTGAGGTCGGGTTCACACCTACTACCTGTATGCTCTAC        1320

Qy  1321  acctcagctcactgcacctctcgcctccccactcaagcaatcctctgttctcagcctcc  1380
Db  1321  ACCTGAGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCAATTCCTGTCTCAGCCTCC    1380

Qy  1381  cgcgtagctgggactacaggcgcacgcccggcctaatttttgtattgttagtagagatggg 1440
Db  1381  CCGGTAGCTGGGACTACAGGCGCCCACCACCGCCTAATTTTTGTATTGTTAGTAGAGATGGG 1440

Qy  1441  gtttcaccatattagccaggctggtctggaactctcagtgatcccaccctc          1500
Db  1441  GTTTCACCATATTAGCCCAGGCTGGTCTCGAACTCCTGACCTCAGTGATCCACCACCTC   1500

Qy  1501  agcctctaaagtgctgggattacaggcgtgagcccaccgcgcccggcc              1560
Db  1501  AGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCGCGCCCGGCC              1560

Qy  1561  ttaataaggaataacttgaatggttactcaaccacacaggaaacagacaaaagctgtga  1620
Db  1561  TTAATAAGGAATAACTTGAATGGTTTACTCAACCACACAGGAAACAGACAAAAGCTGTGA  1620

Qy  1621  taattcaggggatggggaatggtgcaatggttgccatgtgcccagac              1680
Db  1621  TAATTTCAGGGATGGGGAATGGTGCCATGAGCTCGCCTCCGTAGTCCCAGAC          1680

Qy  1681  cactcgttcctcatcactttctccccatcctccattcattat                    1740
Db  1681  CACTCGGTCCTCATCACTTTCTTCCCCATCCTCCATTTATT                     1740

Qy  1741  caccatgtcttttgtggtaagcctccactcgttactcgagtatacatacag          1800
Db  1741  CACCATGCTTTTGTGGTTAAGCCTCCACTCGTTACTGGAAATAAGAGTATACATAAACTAG 1800

Qy  1801  ttccatttgggccatctgtcgtgtctgtatatggggagggacttcct              1860
Db  1801  TTCCATTTGGGGCCCATCTGTCGTGTCTGTATAGGGGAGGGACTTCCT              1860

Qy  1861  tgaagccccgcgaagggttcctctcccagctggggtggcccatacccccaggtgggtcc   1920
Db  1861  TGAAGCCCCGCGGAGGTTTCCTCTCCCAGCTGGGGAGCCCTGCAAGCACCCCAGACTCCT  1920

Qy  1921  tgggtgtcctcggcgagcaacctgcccgtgccactgtcttgttgttcatccactctctcagg 1980
Db  1921  TGGGTGTCCTCGGAGCAACCTGCCCGTGTTTCTCGTGTGTGTTCTCGTGGATCCTCTCTAGG 1980

Qy  1981  gacctgttgcttctattctgtgactcgttcattcattgacaatt                  2040
Db  1981  GGATGTTGAGAGGGGAAGGAGGGGCAGAGCTGAGCCACAGGGGGAGGTGGAGGG        2040
```

```
Db  1981  GACCTGTTGCTTTCTATTTCTGTGACTCGTTCATTCCAGGCATTCATTGACAATT       2040
Qy  2041  tattgagtacttatatctgccagacaccagaagacaaaatgttgagcaaagcagtcactgc 2100
Db  2041  TATTTGAGTACTTATATCTGCCACACCAGACACCAGAAGACAAAATGGTGAGCAAAGCAGTCACTGC 2100

Qy  2101  cctacctctggagtgttgctggagacagtttccatggaagggctcaggaagaaatcaatagcca 2160
Db  2101  CCTTACCTTCGTCTGGGAGTGGTGACAGTTTCTCATGGAAGACGTGCAGGAAGAAAATTTAATAGCCA 2160

Qy  2161  gccaacttaaaccagtcctgtgaaagagaaagtaacaccatcttgaagtagaagaattgtcgcgc 2220
Db  2161  GCCAAACTTAAAACCCAGTGCTGAAAGAGAAAGTAACACCATCTTGAAAGAATTGTCGCGC 2220

Qy  2221  agcaatccttaacacaaggccacctccccagagcccctgtcccatctgcccccgggag     2280
Db  2221  AGCATCCCTTAACGACGCCACCTCCCTTAACGGCCCCCTTCCCATCGTCCCCGGGAGG     2280

Qy  2281  cccccaagcccagggaaaggggatctctccaggtcaggtccctcccatctgagctctgaacgttagtc 2340
Db  2281  CCCCCAAGCCCAAGCCCGAGCTTCAACGGCCTCCGCCACGTCACACGCCTGCCAGCTGGCCT 2340

Qy  2341  gcctcgcttcccgtgaatccgtcctggtgcatctgagactcctggctcccaaggct      2400
Db  2341  GCCTCGCTTCCCGTGAATCGTCCTGGTGCATCTGAGACTCCTTGGGCTCCAGGCT        2400

Qy  2401  ccagaaaggaaaatggagaagggaaaactagtctaacggagactggagaagtgttcc      2460
Db  2401  CCAGAAAGGAAAATGGAGAAGGGAAAACTAGTCTAACGGAGACTGGAGAAGTGTTTC      2460

Qy  2461  ctccagagggaaaggggcctccacgtccaggagaagagaaattccaaggcctcagcagttc 2520
Db  2461  CTTCAGAGGGAAAGGGGCCTTCCACGTCCACCTCCATCAGTCAAGTTTCCAGGATCGCAGTTTC 2520

Qy  2521  tgggggacctggggcctgaaagccaaggctgaa                             2580
Db  2521  TGGGGGACCTGGGGCTGCAGGAAGGGCAAGGCTGA                           2580

Qy  2581  gctgccccagatgttcagttgttcacgggggctgggagttcctgagc               2640
Db  2581  GCTGCCCAGATGTTCAGTTGGTTCAGTTGGAGTTTTCACGGAGTTGGGTTCCGGTGGAGC   2640

Qy  2641  ctttttatctttttctctgctttggaggagagaagtctattcatgaaggatggcagttc   2700
Db  2641  CTTTTTATCTTTTTCTCTGCTTTGGAGGAGAGAAGTCTATTTCATGAAGGATGGCAGTTTC 2700

Qy  2701  ataaagtcagctgttaaaattccagggtgtgtgcatggggtgtgaaaggccttat       2760
Db  2701  ATAAAGTCAGCTGTTAAAATTCCAGGGTGTGTGCATGGGTGTTTCTTCACGGAAGCCTTTAT 2760

Qy  2761  ttaatggggatataggaagaggagtcaattctcctaggagacgttaattcacggaagaTGAC 2820
Db  2761  TTAATGGGAATATAGGAAGCGGACTCATTTCCTAGGCGGTTAATTCACGGAAGAGTGAC   2820

Qy  2821  tgaagtctcttttctctggacactcagccctgtgttgggactacttta             2880
Db  2821  TGGAGTCTTTTTCATGTCTTGGGCCAACTACTCAGCCCTGTGTGGGAACTTTA        2880

Qy  2881  tgcaagacggtcgaaaacctttggaatctcaggagaactccggccatt              2940
Db  2881  TGCAAGAACGGTCGAAAACCTTGGAATCTCGGTTTCTTTTCTTCGGTTCGCCATT      2940

Qy  2941  ggttggctgcgacccgtggaagtgtctctcatcaggagactcgtttctt             3000
Db  2941  GGTTGGCTGCGACCCGTGGCAAGTGTCTCTCTCTCCTGGGAGTTTCTTTCTGGGCT      3000

Qy  3001  ataaagacccttcagcctctgtgtgacactccccgtggattctctaggg            3060
Db  3001  ATAAAGACCCTTCGCGTGTTTCGTGTGTCTGAACACTTCCCTGATTCTCTGATTCG     3060

Qy  3061  ggatgttgagaggggaaggagggggcagagctgagccacagggggag              3120
Db  3061  GGATGTTGAGAGGGGAAGGAGGGGCAGAGCTGAGCCACAGGGGGAGGTGGAGGG       3120
```

Qy   3121 ggacaggaagggcaggcagaaagctcgggtgtgctccaatcagtcctccactcactcagtcagactc 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GGACAGGAAGGGCAGGCAGGCAGAAAGCTCGGGTGTTCCTCCATCAGTCACTGATCAGACTC 3180

Qy   3181 caggaccgagagccagcccacaatgcttccaggaaagcccacaacagccacattttcct 3240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 CAGGACCGAGAGCCAGCCCACAATGCTTCAGGAAAGCTCAACAACAGCCACATTTTCCT 3240

Qy   3241 tccctaagacatagagacatggcatttgccaataacaaatgcaatgaactggt 3300
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 TCCCTAAGCCATAACGACAATGGCATTTGCCAATAACCAACAGAGAATAAACTGGT 3300

Qy   3301 ggtagcttgcctggcatcaagaaagcaagcaagtcagagatg 3360
           ||||||||||||||||||||||||||||||||||||||||
Db   3301 GGTAGCTTTTGCCTGGCATTCAAAAAACTGGGCCAGCAAGTGG 3360

Qy   3361 ttaaacttttcaccctgaccacccacccagtcagtctgacacagg 3420
           ||||||||||||||||||||||||||||||||||||||||||
Db   3361 TTAAACTTTTCACCCTGACCACGCACGCAGTCAGTCTGACACGACTAACTGGT 3420

Qy   3421 agtagacctgcagcgcagggagaggggagatagtgtatgcaagaag 3480
           ||||||||||||||||||||||||||||||||||||||||||||
Db   3421 AGTAGACCTGCAGCGCAGGGGAGGAGGAGAAAAAATGTTCAAGGGAGAAAG 3480

Qy   3481 acagattcattcaaggggtggaattgacccacccagttgtaactggtt 3540
           |||||||||||||||||||||||||||||||||||||||||||
Db   3481 ACAGATTCATTCAAGGGGCTGTAAGATTACTTAGTTCCTTATTAGACAAGATTTG 3540

Qy   3541 gtccagttagcacgcacaagggcaatacccgttcttcttttaacagg 3600
           |||||||||||||||||||||||||||||||||||||||||||
Db   3541 GGAGTTAGCACGCACAAGGGCAATCCCGTTCTTTTAACAGGAGAGAAAACATTCCTAAGGAG 3600

Qy   3601 cctggattctaatactatattttcctttacaagctgagtaatctcgt 3660
           ||||||||||||||||||||||||||||||||||||||||
Db   3601 CCTGGATTTCTAATACTATATTTTCCTTTACAAGCTGAGTAATCTCACAAG 3660

Qy   3661 gtagtaactgaagggctgtaagattactaagttccttattagagtccggg 3720
           |||||||||||||||||||||||||||||||||||||||
Db   3661 GTAGTAACTGAGGGGTTGTAAGATTACTTAGTTCCTTATTAGACTAGTACCCTTGGCTG 3720

Qy   3721 ggagttagcacgcacaagggcaatacccgttcttcttttaacagg 3780
           |||||||||||||||||||||||||||||||||||||||
Db   3721 GGAGTTAGCACGCACAAGGGCAATCCCGTTCTTTTAACAGGAGAGAAAACATTCCTAAGGAG 3780

Qy   3781 taaagccaaacaagaattcaagtcttgctctgactatatgatt 3840
           |||||||||||||||||||||||||||||||||||||||
Db   3781 TAAAGCCAAACAGAATTCAAGCCTAGGTCTTGCTGACTATATGATTGGGTTTTTTGAAAAAT 3840

Qy   3841 cattccggaatgttcactcctcgatcagaaaatgagactagtacc 3900
           |||||||||||||||||||||||||||||||||||||||||
Db   3841 CATTTCAGGGATGTTTACTATCCGATTCAGAAAATGAGACTAGTACCCTTTGGTCAAGCTG 3900

Qy   3901 taaacaacaaacccagttgtaaatgtctcaagcttaactgcagaaccaatcaaa 3960
           ||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TAAACAAACACCCAGTTGTAAATGTCTCAAGTTAACTGCAGAACCAATCAAA 3960

Qy   3961 agaatagaatctttagacaaagctgcaaggc 4020
           |||||||||||||||||||||||||||||
Db   3961 AGAATAGAATCTTTAGAGCAAACTGTGTTCTCCACACTCTGCCAGGGC 4020

Qy   4021 agtttggaaatattttacttcacacactgtgtggtattaacaacataaagt 4080
           ||||||||||||||||||||||||||||||||||||||||||
Db   4021 AGTTTGGAAATATTTTACTTCACACAGTTATTGTTGGTATTAACACATAAAAGT 4080

Qy   4081 tgctcaagcaatcaattatttcaagttggcttaagttacttctgacacagttggtatat 4140
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 TGCTCAAAGGCAATCATTATTTCAAGTGGCTTAAAGTTACTTCTGACAGTTTGGTATAT 4140

Qy   4141 ttattggctattgccattgctttttgttttgttttattaaatgtaaaagcag 4200
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4141 TTATTGGCTATTGCCATTTGCTTTTTGTTTTTGTTTATTAAATGTAAAAGCAG 4200

Qy   4201 ggattattaacctacagtccagaagcctgtgaattgaatgaggaaaattacattt 4260
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGATTATTAACCTACAGTCCAGAAAGCCTGTGAATTGAAATGAGGAAAATTACATTTT 4260

Qy   4261 tgtttttacccctcctaactaaaattattattattattccattgggaattagagcatgaa 4320
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4261 TGTTTTTACCACCTTCCTAACTAAATTTAACATTTTATTCCATTGGGAATTAGAGCCATAAA 4320

Qy   4321 ctcaaagtggtaataacagtacctgtgatcttgtcattaccaatagaaatcacagacatt 4380
           |||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CTCAAAGTGGTAATAACAGTACCTGTGATTTTGTCATTACCAATAGAAATCACAGACATT 4380

Qy   4381 ttatacttattacagttgttgcagataagtgtaaagtgaaatactcaaaacta 4440
           ||||||||||||||||||||||||||||||||||||||||||
Db   4381 TTATACTATTACAGTTGTTGCAGATAAGTACCTCTGATTTGTTGCATTAACCAATGTACTCAAAAACTA 4440

Qy   4441 ctttgaaattagacctcctgctgttcttgtttttaaacatattaataataaacatgttaaa 4500
           |||||||||||||||||||||||||||||||||||||||||||
Db   4441 CTTTGAAATTAGGACCTCCTGCTGTTGTTTTTTAAACATATTAATAATAAACATGTTAAA 4500

Qy   4501 atttgatatttgataatcattcattgtttccttgtaatctatatttt 4560
           |||||||||||||||||||||||||||||||||||||||||||
Db   4501 ATTTTGATATTTGATAATCATTCATTTGTATATCTATATTTT 4560

Qy   4561 atatatttgaaaaaccatctttctgggaagagaagtcccagaaccagtggttccttgg 4620
           |||||||||||||||||||||||||||||||||||||||||||||
Db   4561 ATATATTTGAAAAACATCTTTCTAGGAATGCTCTCCCCAGTGAGCCGTGGTTCTTGG 4620

Qy   4621 catgccaacacagagtaagaaactgatttagagactacttgacaattggttgcctgaga 4680
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 CATGCACACAGAGTAAGAACGTAAGAAACTGATTTAGAGGCTAACATTGGTCGCCTGAGA 4680

Qy   4681 tgcaagactgaaattagaaagttctccccaaagtacacagttgtttccttgtaatctcatttt 4740
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 TGCAAGACTGAAATTAGAAAGTTCTCCCAAAGATACACAGTTGTTTAAAGCTAATCTATATTTT 4740

Qy   4741 aggggggaaatctgccgcttcctataggaagtgctctcccaggaagttgttgctgtc 4800
           |||||||||||||||||||||||||||||||||||||||||||
Db   4741 AGGGGGGAAATCTGCCGCTTCCTATAGGAAGTGCTCTCCCAGGAAGTTGTTGGTGTC 4800

Qy   4801 cttgtgttctggctggttgttttttctcctgttcctgctcacgtcttcagcttaagggactt 4860
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CTTTGTGTTCTTGGCGGCTGGTTGTTTTTTCTCCTGTTCCTGCTCCACGGTCTTAAAGGACTTGTTT 4860

Qy   4861 ggatctccagtcctagcatagtgcctggcacgtcctcaatgagtttgcagag 4920
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4861 GGATCTCCAGTCCTAGCATAGTGCCTGGCACAGTTCTCAATGAGTTTTCCAGAG 4920

Qy   4921 tgaatggaaatataaactagaaaaatatatcccttgtgaaatcagcagacgacacagtagtcctgg 4980
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TGAATGGAAATATAAACTAGAAAAATATCCTTGTTGAAATCAGCAGACACCAGTAGTCCTGG 4980

Qy   4981 tgttaagtgtgttacgtgtgtgttgtgttgtgttgtgtggtgggata 5040
           |||||||||||||||||||||||||||||||||||||||||||
Db   4981 TGTAAGTGTGTTACGTGTGTGTGTGTGTGTGTGTGGTGGGGGATA 5040

Qy   5041 taggaactattattggggtatggtgcataaaattgggatgtctttgaaaaagagaaactc 5100
           ||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TAGGAACTATTATTGGGGTATGGTGGCATAAAATTTCTTTTTAAAAGGAAACTC 5100

Qy   5101 caaacagacttctctgggaaggttattttctaagaaatcttgctgctggcagcgtgaaggcaacccc 5160
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 CAAACAGACTTCTCTGGGAAGGTTATTTTCTAAGAAATCTTGCTGCCAGCGTGAAGGCAACCCC 5160

Qy   5161 cctgtgcacagcgccccaacgtggccacctcgtctctcccccatgatgacgaggggctg 5220
           |||||||||||||||||||||||||||||||||||||||||||||
Db   5161 CCTGTGCACAGCGCCCCAGCGTGGCCACCTCGTCTCTCCCCCATGAGCACGAGGGGCTG 5220

Qy   5221 gctccccagtatataaaacctcctggagctcgggcatgagccagccaagg 5271
           |||||||||||||||||||||||||||||||||||||||||||
Db   5221 GCTCCCCAGTATATAAAACCTCTCTGGAGCTCGGGCATGAGCCAGCCAAGG 5271

```
RESULT  2
LOCUS       HS454G6/c
LOCUS       HS454G6        79376 bp    DNA             PRI       23-NOV-1999
DEFINITION  Human DNA sequence from PAC 454G6 on chromosome 1q24. Contains
            trabecular meshwork inducible glucocorticoid response protein,
            TIGR, myocilin, ESTs and STS.
ACCESSION   Z98750
VERSION     Z98750.1  GI:2887277
KEYWORDS    1q24; myocilin; TIGR.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79376)
  AUTHORS   Deadman,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1997) Chromosome 1 Project Group, Sanger Centre, Hinxton,
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            (http://www.sanger.ac.uk/HGP/Chr1/) Sanger Centre, Hinxton,
COMMENT     On Feb 14, 1998 this sequence version replaced gi:2465060.
            IMPORTANT: This sequence is not the entire insert of clone 454G6.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations.
            together with a note of the overlapping clone name. Note that the
            variations annotated may not be found in the sequence submission
            corresponding to the overlapping clone as we submit sequences with
            only a small overlap as described above.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 1, constructed by the Sanger Centre chromosome 1
            mapping group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1/
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The true left end of clone 560B9 is at 79273.
            The true left end of clone 454G6 is at 1 in this sequence. The true
            454G6 is from the library RPCI3 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong.
            For further details see http://bacpac.med.buffalo.edu/.
FEATURES             Location/Qualifiers
     source          1..79376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q24"
                     /clone="RP3-454G6"
                     /clone_lib="RPCI-3"
     prim_transcript 1914..>3968
                     /note="19 copies of 2 mer 82 % conserved"
     repeat_region   435..472
                     /note="22 copies of 2 mer 89 % conserved"
     repeat_region   3703..3746
     match: N30491 AA307150 AA192"
     match: H08333 H08236 N42052 D61944 R27102 N32353
     match: AA046487 AA369741 H08313 AA168895 H32730
     match: AA329084 W47082 AA043955 AA341783 AA353681
     match: AA313383 AA163561 N89173 AA174814 AA057059
     match: AA163561 F02925 AA131540 W00634 R36066
     match: F02925 AA131540 W00634 R36066 AA313383
     match: R56676 AA043968 W63639 F12081 AA046699
                     /note="match: multiple ESTs
     repeat_region   4051..4183
                     /note="AluSq repeat: matches 1. .133 of consensus
     repeat_region   4200..4502
                     /note="incomplete repeat"
     repeat_region            /note="AluSq repeat: matches 2. .301 of consensus"
```

```
     repeat_region   4659..4851
                     /note="AluSx repeat: matches 2. .194 of consensus
     repeat_region   5216..5345
                     incomplete repeat"
                     /note="AluJo repeat: matches 132. .1 of consensus
     repeat_region   7759..7907
                     /note="MIR repeat: matches 174. .1 of consensus"
     repeat_region   7933..9328
                     /note="TIGGER1 repeat: matches 1. .1472 of consensus"
     repeat_region   9332..9626
                     /note="AluSg repeat: matches 1. .289 of consensus"
     repeat_region   9639..10335
                     /note="TIGGER1 repeat: matches 1469. .2174 of consensus"
     misc_feature    <10903..>16855
                     /note="endogenous retroviral sequence"
     repeat_region   10343..10642
                     /note="TIGGER1 repeat: matches 2175. .2417 of consensus"
     repeat_region   10643..10856
                     /note="AluSg repeat: matches 1. .300 of consensus
     repeat_region   11049..11384
                     /note="LTR2 repeat: matches 31. .449 of consensus"
     repeat_region   15835..15870
                     /note="18 copies of 2 mer 83 % conserved"
     repeat_region   16856..17286
                     /note="AluY repeat: matches 301. .2 of consensus"
     repeat_region   17287..17574
                     /note="LTR2 repeat: matches 31. .449 of consensus"
     repeat_region   18294..18650
                     /note="AluJb repeat: matches 15. .300 of consensus"
     repeat_region   18877..19180
                     /note="THE1B repeat: matches 358. .1 of consensus"
     repeat_region   19767..20013
                     /note="AluY repeat: matches 301. .2 of consensus"
     repeat_region   20264..20722
                     incomplete repeat"
                     /note="AluJb repeat: matches 131. .1 of consensus
     repeat_region   20858..21223
     consensus"
                     /note="MLT1F repeat: matches 482. .1 of consensus"
     repeat_region   21216..21302
     consensus"
                     /note="MSTD repeat: matches 394. .1 of consensus"
     repeat_region   21403..21703
     consensus"
                     /note="MLT2_internal repeat: matches 4520. .4433 of
     repeat_region   21978..22357
     consensus"
                     /note="MLT2_internal repeat: matches 5002. .4750 of
     repeat_region   22051..22118
     consensus"
                     /note="MLT2_internal repeat: matches 3887. .3580 of
     repeat_region   22363..22524
     consensus"
                     /note="MLT2_internal repeat: matches 2495. .2317 of
     misc_feature    22286..>23680
                     complement(22652..24072)
                     /note="STS G07544"
     misc_feature    22286..>23680
                     /note="STS G07436"
     repeat_region   23007..23309
                     /note="AluSp repeat: matches 3. .301 of consensus"
     repeat_region   22551..22839
                     /note="AluSx repeat: matches 302. .1 of consensus"
     repeat_region   24008..24290
                     /note="AluJo repeat: matches 298. .6 of consensus"
     repeat_region   25895..26364
                     /note="LTR2 repeat: matches 2. .449 of consensus"
     repeat_region   26397..26697
                     /note="AluSp repeat: matches 1. .302 of consensus"
     repeat_region   26719..27021
                     /note="AluSq repeat: matches 1. .302 of consensus"
     repeat_region   29037..29334
                     /note="AluSx repeat: matches 1. .302 of consensus"
```

```
repeat_region   /note="AluSx repeat: matches 300..3 of consensus"
30028..30310
repeat_region   /note="AluSx repeat: matches 1..301 of consensus"
30402..30539
repeat_region   /note="MLT2B repeat: matches 264..404 of consensus"
30402..30615
prim_transcript /note="MLT2A repeat: matches 264..453 of consensus"
<30801..>31136
                /note="match: 5' EST AA256902 clone 682136"
repeat_region   /note="MER4B repeat: matches 199..431 of consensus"
30980..31222
repeat_region   /note="MER4B repeat: matches 417..540 of consensus"
31232..31528
repeat_region   /note="AluSc repeat: matches 299..1 of consensus"
31530..31654
repeat_region   31836..32135
repeat_region   /note="AluY repeat: matches 300..1 of consensus"
32200..32301
repeat_region   /note="AluJo repeat: matches 186..287 of consensus
                incomplete repeat"
32365..32493
repeat_region   /note="AluY repeat: matches 300..1 of consensus"
33283..33399
repeat_region   /note="MIR repeat: matches 206..67 of consensus"
33408..33667
repeat_region   /note="AluJ repeat: matches 77..196 of consensus"
33955..3043
repeat_region   /note="AluSc repeat: matches 1..291 of consensus"
33955..3043
repeat_region   /note="MIR repeat: matches 80..167 of consensus"
34725..34904
repeat_region   /note="AluJo repeat: matches 299..127 of consensus
                incomplete repeat"
34907..35207
repeat_region   /note="AluY repeat: matches 300..1 of consensus"
35212..35344
repeat_region   /note="AluJb repeat: matches 133..1 of consensus
                incomplete repeat"
35545..36842
repeat_region   /note="AluJo repeat: matches 299..1 of consensus"
38190..38379
repeat_region   /note="MER3 repeat: matches 209..13 of consensus"
38382..38682
repeat_region   /note="AluSx repeat: matches 1..302 of consensus"
38836..38968
repeat_region   /note="MIR2 repeat: matches 145..2 of consensus"
38983..39279
repeat_region   /note="AluJb repeat: matches 3..296 of consensus"
39418..39773
repeat_region   /note="THE1A repeat: matches 354..2 of consensus"
39775..40120
repeat_region   /note="THE1B-INTERNAL repeat: matches 1580..1234 of
                consensus"
40757..41023
repeat_region   /note="AluJo repeat: matches 37..301 of consensus
                incomplete repeat"
41307..41589
repeat_region   /note="AluSq repeat: matches 21..303 of consensus
```

```
Query Match     99.3%;   Score 5232.4;   DB 65;   Length 79376;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 5267;  Conservative   0;  Mismatches   1;  Indels   4;  Gaps   3;
```

```
Qy       1 atctttgttcagtttacctcaagggctatatgaagtgaagtggataaaccaatgtgaaag 60
           ||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||
Db   68406 ATCTTTGTTCAGTTTACCTCAGGGGTTATTATGAAGTGAAGTGGATAAACCAATGTGAAAG 68347

Qy      61 tcctatataaactgtatagcctccatccggatgtatgtgtcttggcaggatgataagaatca 120
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68346 TCCTATAAACTGTATAGCCTCCATCCGGATGTATGTGTCTTGGCAGGATGATAAGAATCA 68287

Qy     121 ggaagaaggagtatccaacgttagccaacagttatcccaagtgtgtcgctcttatttagtga 180
           ||| |||||||||||| |||||||||||| |||||||||| |||||| ||||||||||||
Db   68286 GGAGGAAGGAGTATCCACGTTAGCCAACGTTGTCCAGGCTGTGTCTCCTCTTATTTTAGTGA 68227

Qy     181 cagatgttgtcctgacagagagcttattcttcaggaaacacatcaccaaatggtaaatc 240
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68226 CAGATGTTGCTCCTGACAGAGAGCTTATTCTTCAGGAAACACATCACCAAATGGTAAATC 68167

Qy     241 catcaaacaggcagctaagaaacacaggaacacgagaaacttgcccaaggagaaaaatgccag 300
           ||||||||||||||||||||||||||||||         ||||||||||||||||||||||
Db   68166 CATCAAACAGGCAGCTAAGAAACACAGGAATGAGATGGGCACTTGCCCAAGGAGAAAAATGCCAG 68107

Qy     301 gagagcaaaataatgatgaaaaataaactttccccttttgttttaatttcaggaaatg 360
           |||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||||
Db   68106 GAGAGCAAAATAATGATGAAAAATAAACTTTCCCTTTGTTTTAATTTCAGGAAATG 68047

Qy     361 atgaggacaaaataatgaataagaaaacagctcaagaaaacaagagtgttctccaaatctc 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68046 ATGAGGACAAAATAATGAGAAAAGAAAATGGCTCAGGAAAACAAGAGTGTTCCCAAATTGG 67987

Qy     421 taattaagtattgttccttgggaagaagagaacctccatgtgagcttgatgtgggaaaactca 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67986 TAATTAAGTATTGTTCCTTGGGAAGAAGAGAACCTCCATGTGAGCTTGATGTGGGAAACTC 67927

Qy     481 aaacgtcaaaagcatgatctgatccaaagtggatattattattttaaaacccagat 540
           ||| ||||||||||||||||||||| ||||||||||||||||| ||||||||||||||
Db   67926 AAACGTCAAAAGCATGATCTGATCCAGAGTGGATATTATTATTTTAAAAACCAGAT 67867

Qy     541 ggcatcaactctgggagggcaagtcaggtcatgttcgttgcaaggacactaacaataac 600
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   67866 GGCATCACTCTGGGAGGGCAAGTCAGGTCATGTTCGTTGCAAGGACACTAACAATAAC 67807

Qy     601 agcaaaatcaaaatccgcaaaatgcagaggagcagaaaatgggaactgggaaagctttcatac 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67806 AGCAAAATCAAAAATCAAAATCCGCAAAATGCAGAGGAGCAGAAAATGGGACTGGGAAAGCTTTCATAAC 67747

Qy     661 agtgattaggcagttgaccatgttccgtcaacacacctcccgtctataccaggagaacacaaaa 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67746 AGTGATTAGGCAGTTGACCATGTTCCGTCAACACCTCCCGTCTATACCAGGAGAACACAAA 67687

Qy     721 attgactgggctaagccctggactttcagagaaatatgagaacgacaaaaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67686 ATTGACTGGGCTAAGCCCTGGGACTTTCAAGAGAAATATGAGAGCGACAAAAA 67627

Qy     781 gacatcgttaaaagtgcaaccacagtcatgttgagcctttcaaggcagccagtgcccctagca 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67626 GACATCGTTAAAAGTGCAACCACAGAGACATTGTGAGGCTTCAAAGCACCAGTGCCCCTCAGCA 67567

Qy     841 gggaccctgaggcattgcctttaggaaggaataggccagttcttaggagaatcttaagagaact 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67566 GGGACCCTGAGGCATTTGCCTTTAGGAAGGCCAGTTTCTTAGGAGAATCTTAAGAGAACTC 67507

Qy     901 ttgaaagatcaatgaatcttaaccatttttaagtataaaacaaatatgcgatcagcataatcag 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67506 TTGAAAGATCAATGAATCTTAACCATTTTTAAGTATATAAACAAATATGCGATCAGCATAATCAG 67447

Qy     961 tttagacatggttcccaaatttttataaagtcaaggcatacaagtgtcccagctcc 1020
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   67446 TTTAGACATGGGTCCCAAATTTTTATAAAGTCAAGGCATACAAGTGTCCCAGCTCC 67387

Qy    1021 ggatagtgtcagaaatccattagaaatcaactgtgtccccacatcctaactttttcagaatgatc 1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67386 GGATAGTGTCAGAAATCATTAGAAATCACTGTGTCCCCACTACAACCTTTTTCAGAATGATC 67327

Qy    1081 tgtcatagccctcacacaggcccgatgtgtgtctgaccctaacacaacaacctcccc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67326 TGTCATAGCCCTCACACAGGCCCTCACACACAGGCCCGATGTGTCTGACCCTACAACCACCCAA 67267

Qy    1141 gtgcctcaaccattgttaacgtgtcatccagtaggtccccactacaaatgccacactcccc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   67266 GTGCCTCAACCATTGTTAACGTGTCATCCCAGTAGGTCCCCACTACAAATGCCACACTCCCC 67207

Qy    1201 ggaagaaggagtctcccccacttctgacctcacggatgt 1260
           ||||||||||||||||||||||||||||||||||||||
Db   67206 TGTTGCAGCCCATCCCGCTCCACAGGAAGTCTCCCCCACTTCTGACTTCTGCATCACGATGT 67147

Qy    1261 tacagccgagagctccgtgggggtgaggtctgtgtcttacacctacctgtatgctctac 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68286 GGAGCAAAGGAGTATCCACGTTAGCCAAGTGTCCAGGCTGTGTCTCCTCTTATTTTAGTGA 68227
```

```
Db  67146  TACAGCCAGAAGCTCCGTGAGGGTGAGGGTCTGTGTCTTACACCTACTGTATGCTCTAC  67087
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1321  acctgagctcactgacgttcaagtcaattctcctgtctccagcctcc              1380

Db  67086  ACCTGAGCTCACTGACGTTCAAGTCAATTCTCCTGTCTCCAGCCTCC              67027
           |||||||||||||||||||||||||||||||||||||||||||||||
Qy   1381  cgcgtagctgggactacaggcgcacgccaccacgcccggctaattttgtattttagtagg 1440

Db  67026  CGCGTAGCTGGGACTACAGGCGCACGCCACCACGCCCGGCTAATTTTTGTATTGTTAGTAGAGATGGG  66967
           |||||||||||||||||||||||||||||||||||||||||||
Qy   1441  gtttcaccatattagggccggctggtctcgaacctcctgacctcaccacctc         1500

Db  66966  GTTTCACCATATTAGCCCGGGTTGTCTTGAACCTCTGACCTCACCACCTC           66907
           ||||||||||||||||||||||||||||||||||||||||||||||
Qy   1501  agccctctaagtgtcgggaatacagcatgggtcagtgt                       1560

Db  66906  AGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAGTGT         66847
           ||||||||||||||||||||||||||||||||||
Qy   1561  ttaataaggaataacttgaatggttactaaaccaacaggaaacacagctgtga         1620

Db  66846  TTAATAAGGAATAAACTTGAATGGTTTACTAAACCAACAGGAAACGCTGTGA          66787
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1621  taattttccattgggggattcttgggaatgggggaggg                       1680

Db  66786  TAATTTCAGGGATTCTTGGGAATGGGTTGTCTTGGAGGTGCCATGG                66727
           |||||||||||||||||||||||||||||||
Qy   1681  cactgttcctcatcacttctcctcatcccatttcaggcta                     1740

Db  66726  CACTGGTCCTCACACTTTCTTCCCTCATTTCAGGCTAAGTTACC                  66667
           ||||||||||||||||||||||||||||||
Qy   1741  caccatgcttttgttggtaagctcccacatcgttactgaaatac                 1800

Db  66666  CACCATCCTTTTGTTGGTAAGCTTACTAACAACAGGGAAAACAGCTGA              66607
           |||||||||||||||||||||||||
Qy   1801  ttccattcggggttcctgtgtggtgtatggggg catcca                     1860

Db  66606  TTCCATTTGGGGGCCATCTGTGTGGTGTATAGGGGAGGGGCCATAC                66547
           ||||||||||||||||||||||||||||||
Qy   1861  tgaagtcccgccagaggtttctcc                                     1920

Db  66546  TGAAGCCCCGGCCAGAGTTTTCTCTCGTGCGGACCGTGGGC                     66487
           |||||||||||||||||||||||
Qy   1921  tgggtgtcctgagcaacctgccagccagcccgtccactgttgt                  1980

Db  66486  TGGGTGTCCTGAGCAACCTGCCAGCCCGGTCCCACTGGTT                      66427
           |||||||||||||||||||||||||||||
Qy   1981  gacctgttgcttctcttcgtgttcca                                   2040

Db  66426  GACCTGTTGCTTCTTCTCGTTCATTCAGG                                 66367
           |||||||||||||||||||
Qy   2041  tattggtacttatatctgccagacaacagagacaaaatgttga                  2100

Db  66366  TATTGAGTACTTATATCTGCCAGACAGACAAAATTGGTCAGC                    66307
           |||||||||||||||||||||||||
Qy   2101  cctacccttcgtggagtggacagtttcctcc                              2160

Db  66306  CCTACCCTTCGTGGAGGTGCAGAGAGTTTCATGGAAGA                        66247
           |||||||||||||||||||||
Qy   2161  gccaacttaaacccagtgctgaaagaaaaatcccacc                        2220

Db  66246  GCCAACTTAAACCCAGTGCTGAAAGAAAAATAAACACCA                       66187
           |||||||||||||||||||||||||||||||||
Qy   2221  agcatccccttaacaagcgcaacctctg                                 2280

Db  66186  AGCATCCCTTAACAAGGCCACCTCGTGCC                                 66127
           ||||||||||||||||||
Qy   2281  cccccaagccgagtcttccaagcctcc                                  2340

Db  66126  CCCCCAAGCCCAGTCTTCCAAGCCTCTCC                                 66067
           ||||||||||||||||||
Qy   2341  gcttcgcttccccgtgaatcgtcggt                                   2400
```

```
Db  66066  GCCTCGCTTCCCCTGAATCGTCGGTCCATCTGGCTGGGACTCCTTGGGGCT           66007
           |||||||||||||||||||||||||||||||||||||
Qy   2401  ccagaaaggaaatggagagaggg aaaactagtctaac gttctcc               2460

Db  66006  CCAGAAAGGAAATCGAGAGAGGG AAAACTAGTCTAACGGAAATCTGTTTC           65947
           ||||||||||||||||||||||||||||||||
Qy   2461  ctccagagaggggaggg cctccaagtccaggggg                          2520

Db  65946  CTCAGAGGGGAAAGGGGGCCTTCCACCTCAGGGGAGGGGAACTCAGGGGAG           65887
           |||||||||||||||||||
Qy   2521  tggggacggctggggtctgaagtcc                                    2580

Db  65886  TGGGGACGGCTGGGGTGGTTCTGAAAGGGTGAAAGGCCAAGGCTGAA              65827
           ||||||||||||||||||||||||
Qy   2581  gctgccagatgttcagttgttccacggggtgggg                           2640

Db  65826  GCTGCCCCAGATGTTCAGTTCAGTCGTTGTTCACGGGGAGTGTTCCGTTGAGC         65767
           ||||||||||||||||||||||||||
Qy   2641  ctcagagtggggagggcctccaagtccaggggg                            2700

Db  65766  CTTCAGAGTGGGGATGTTCTTCCATGTCTTGGGAGTGGGGAGTGACTGGCTTA         65707
           ||||||||||||||||||||||
Qy   2701  ataaagtcagctgttaaaactccaggg                                  2760

Db  65706  ATAAAGTCAGCTGTTAAAATTCAGGGGTGCATGGGTTTCCTTCACGAGGGGTC         65647
           ||||||||||||||||||||||||||
Qy   2761  ttaatgggggatataggagagctccaatt                                2820

Db  65706  TTAATGGGGATATAGGAGAGCTCCAATTCCTAGGG                           65647
           |||||||||||||||||||||||||||||
Qy   2821  tgggagtctctttcttctctgtctcctgg                                2880

Db  65766  TGGAGTCTTTATCTTTTCTCATGTCTTCTGGGGAACTCTATTTCAGGG              65707
           |||||||||||||||||||||||||
Qy   2881  tgcaagacggtcgaaaaccttg aatcagga                              2940

Db  65706  TGCAAGACGGTCGAAAATCTTGAATCAGGAGGTTTTCTTTTCTTTCTGCCATT         65647
           |||||||||||||||||||||
Qy   2941  ggttggctgtcgaccgtgggcaagtgtcctc                              3000

Db  65466  GGTTGGCTGTCGACCGTGGGCAAGTGTCTCTTCATTCACGGGGCCGTTAATTC         65407
           ||||||||||||||||||||||||
Qy   3001  ataaagacccttgcagctccgtctgtctc                                3060

Db  65406  ATAAAGACCCTTGCAGCTCTCGTTCTTCTGTCTCGTGTAGGG                    65347
           |||||||||||||||||||||
Qy   3061  ggatgttgaggggtggagcca                                        3120

Db  65346  GGATGTTGAGGGGTGGAGCCAGGAGGGGAAGAGGGG                          65287
           ||||||||||||||||||||||
Qy   3121  ggacaggaaggcaggaaggctgg                                      3180

Db  65286  GGACAGGAAGGCAGGAAGGCTGGGGTGGTCCAAGAGAGCTGGAGG                 65227
           |||||||||||||||||||||||
Qy   3181  caggaccgagagccacaatgtctcag                                   3240

Db  65226  CAGGACCGAGCAGCCACAATGTCTCAGGGAAAGCCCAACAGCCACATTTTCCT         65167
           ||||||||||||||||||||||
Qy   3241  tccctaagcatagacaacaatggcattt                                 3300

Db  65166  TCCCTAAGCATAGACAACAATGGCCAATACCAAAAAGAGAATAACCA              65107
           |||||||||||||||||||||||
Qy   3301  ggtagcttgtcagcattcaaaaactg                                   3360

Db  65106  GGTAGCTTTTGCCTGGCATTCAAAAAACTGGCAAAGTGGAAATGCCAGAGATTG        65047
           |||||||||||||||||||||
Qy   3361  ttaaactttcacccctgacca                                        3420

Db  65046  TTAAACTTTTTCACCCCTGACCAGCAGCACCCCACGCACTGCTGACAGCACGG         64987
           |||||||||||||||||||
Qy   3421  agtgacctgcagcgagggga                                         3480

Db  64986  AGTGAACCTGCAGCGAGGGAGGAGAAGAAAGAGAGGGG                        64927
```

```
Qy    3481 acagattcattcaaggggcggtgggattatagtagttcacctgtgatcctgg 3540
           |||||||||||||||| || ||||||| ||| || ||||||||| |||||||||
Db   64926 ACAGATTCATTCAAGGGACGGTGGGAATTGACCACACGGGATTATAGTCCACGTGATCCTGG 64867

Qy    3541 gttctagggagggcaggggctatattgtggggggaaaaaatcagttcaaggga 3600
Db   64866 GTTCTAGGGAGGGGCAGGGGCTATATTGTGGGGGGAAAAAATCAGTTCAAGGGAGA 64807

Qy    3601 cctgattcttaatacatactatattttttcctttcctttgagcaagtcacaag 3660
Db   64806 CCTGATTCTTAATACTATATTTTTCCTTTCCTTGAGCAAGTCGGGAGA 64747

Qy    3661 gtagtaactgaggtctgtaagatattttttccttattaggaaactcttcctgt 3720
Db   64746 GTAGTAACTGAGGGTGTAAGATTACTTTTCCTTATTAGGAAACTGAGCAAGTCACAAG 64687

Qy    3721 ggagttagcacaaggcaatcccggttcctcttcttttcttttcttccttttctctgt 3780
Db   64686 GGAGTTAGCAGCACAAGGGCAATCCCGTTTCTCTTTTCTTTTCTTCCTTTTTCTCTGT 64627

Qy    3781 taaagcccaaacagattcaggcctggttcctgctctgctcctaagtaaggaact 3840
Db   64626 TAAAGCCAAACAGATTCAAGCCTAGGTTCTTGCTCTGCTCCTAAGGAGACTTGAGAG 64567

Qy    3841 cattcagcgatgtcttactctctgattcagagaaatgaggactagtaccctttggtcagctg 3900
Db   64566 CATTTCAGCAGTGTTTACTANTCTGATTCAGAAAAATGAGGACTAGTACCCTTTGGTCAGCTG 64507

Qy    3901 taaacacaaccaccagttgtaaagtgttcttcaaggcttaactgagaaccaattcaaa- 3959
Db   64506 TAAACAAAACACCCATTGTAAATGTCTCAAGGCTTAACTGCAGGAACCAATCAAAT 64447

Qy    3960 aagaatagaaacatcttggaacaacctggaggtgagtctgccagg 4019
Db   64446 AAGAATAGAAATCTTTAGAGCAAAACTGTGTTTCTCCAC-TCTGGAGGTGAGTCTGCCAGG 64388

Qy    4020 cagtgtttggaaatatattactcacacaagtatggtgtttattaacacataagg 4079
Db   64387 CAGTTGGAAATATTACTTCACAGTAGTTGGTATTAACAACATAAAG 64328

Qy    4080 ttgctcaaggcaatcattattctctgacacagtttggtata 4139
Db   64327 TTGCTCAAAGGCAATCATTATTTCAAGTGGCTTAAAAGTTACTTCTGACACGTTTGGTATA 64268

Qy    4140 tttattgggtcattgtcctttttgttttttctctttctttattaatgtaaagca 4199
Db   64267 TTTATTGGGCTANTGGCCATTTTGTTTTTTCTCTTTCTTTATTGTAANTGAAAGCA 64208

Qy    4200 gggattattaacctacagtccagaaagccctgtgaatttgaatgaggaaaaaatcacatt 4259
Db   64207 GGGATTATTAAACTACAGTCCAGAGAAAGCCCTGTGAATTTGAATGAGGAAAAAAATTACATTT 64148

Qy    4260 ttgtcttcaccctctaactaaaatttattccattgggaattagagtagagccataa 4319
Db   64147 TTGTTTTTACCACCTTCTAACTAAAATTTAACATTTTATTCCATTTGCGAATAGAGCCATAA 64088

Qy    4320 acccaaaggtgttaatacaacgtacctgtgatttgtttgtttttattaaccaatacacagacat 4379
Db   64087 ACTCAAAGTTGGTAATAAACAGTAACCTGTGATTTGTTGTCATTACCAATCACAGACAT 64028

Qy    4380 tttatactatattacagttgttgccagatacgttgtaagtgaaatatttttatcactccaaaact 4439
Db   64027 TTTATACTATATTACAGTGTTGCAGAATTGCTGGTTGTTCTGCAAAAAAAT TACATTT 63968

Qy    4440 acttgaaattagacctcctctggatcttgttttttcattttttaacatataaaaacatgttaa 4499
Db   63967 ACTTTGAAATTAGACCTCTGCTGGATCTTTTTTAACATATTTAAATAAAACAGTTTAA 63908

Qy    4500 aatttgatcattttgatatcattcattgttcctttgtaatctcatatt 4559
Db   63907 AATTTTGATATTTGATAATCATATATTTGTTTCCTTTTGTAANTCATATATTT 63848

Qy    4560 tatatatttgaaaaaacatcttctgaagaggttcaccaatgagagttcactgg 4619
Db   63847 TATATATTTGAAAACANTCTTTCTGAGAAGAGTTCCCCAGAGATTCACCAAATGAGGTTCTTG 63788

Qy    4620 gcatgcaacacacagaagtaagaacttgatttagaggctaacaattgacacttggtgtctgag 4679
Db   63787 GCAATGCAACACACAGAGTAAGAAACTTGATTTAGAGGCTAACATTGACATTGGTCTGCCTGAG 63728

Qy    4680 atgcaagaactagaaaagtttcttcccaaagaatacacagtgtttttaaagctagggt 4739
Db   63727 ATGCAAGACTGAAAATTAGAAAGTTCTCCCAAAGAATACACAGTTGTTTTTAAAAGCTAGGGGT 63668

Qy    4740 gaggggggaaatcttgccggcttcctcatatagggactaggaagttgaggtgctgt 4799
Db   63667 GAGGGGGGAAAATCTCGGCGCTTCCTCATAGGAAGGCTCCCGGGGAGCCTGGTGTAGGGTCGT 63608

Qy    4800 cccttgttcttggctcggctgtgtattctttctctgtgtacgtctttaaaggaacttgtt 4859
Db   63607 CCTTGTTCTTGGCTGGCTGTGTTATTTTTCTCTGTGTACGTCTTTTAAAGGACTTGTT 63548

Qy    4860 tggaatccccagtgttcctagcattggctgttaattttttcttgtgccctgt 4919
Db   63547 TGGATCTCCAGTTCTGGCTGTTAATTTTTCTGTGCCCTGT 63488

Qy    4920 gttgaatggaaatataaacatggaaaaataatcctgttgttcacagtttcctgg 4979
Db   63487 GTGAATGGAAAATATAAACATAGAAAATATATCCCTTGTTGTTCACAGTTCCTG 63428

Qy    4980 gtgtaagtgtgttgtaacgtcgtgtgtgtgtgtgtgaaggtggaget 5039
Db   63427 GTGTAAGTGTGTTAC--GTGTGTGTGTGTGTGTGTGTGTTGTAAAACAGGTGGAGAT 63370

Qy    5040 ataggaactattattgggtatgggtgcataaaatgggatgtttcttttctttttaaaagaaact 5099
Db   63369 ATAGGAACTATTATTGGGGTTGCATAAATGGGATGTTTCTTTTTAAAAGAAACT 63310

Qy    5100 ccaaacagaacttctggaaggttattttcctaagaaatcttgctggcgtgaagggcaacc 5159
Db   63309 CCAAACAGACTTCTGGAAGGGTTATTTCCTAAGAAATCTTGCTGGCGTGAAGGGCAACCC 63250

Qy    5160 ccctgtgcacagccgcccacccagctggccaacctctgtgctcccccatgaagggct 5219
Db   63249 CCCTGTGCACAGCCGCCCACCCAGCTGGCCACCTCTGTGCTCCCCCATGAAGGGCT 63190

Qy    5220 ggctccccagtatatacataaacctctctgaggtcgggcatgagccagcaagg 5271
Db   63189 GGCTCCCCAGTATATACATAAAACCTCTCTGGAGCTCGGGCATGAGCCAGCAAG 63138
```

```
           Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
         * Howland,J.C. Illev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
         * Klein,J., Landers,T., Largcque,K., Lehoczky,J., Levine,R.,
         * Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,P., McCarthy,M.,
         * McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
         * Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
         * Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M.,
         * Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
         * Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
         * Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
         * Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
         * Travers,M., Triglilo,J., Vassilliev,H., Viel,R., Vo,A., Wilson,B.,
         * Wu,X., Wyman,D., Ye,W.J., Young,G., Zalnoun,J., Zimmer,A. and
         * Zody,M.
TITLE    Direct Submission
JOURNAL  Submitted (28-FEB-2000) Whitehead Institute/ MIT Center for Genome
         Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT  On May 25, 2000 this sequence version replaced gi:724945.
         * All repeats were identified using RepeatMasker:
         * Smit,A.F.A. & Green,P. (1996-1997)
         * http://ftp.genome.washington.edu/RM/RepeatMasker.html
         *  ------------------ Genome Center
         *  Center: Whitehead Institute/ MIT Center for Genome Research
         *  Center code: WIBR
         *  Web site: http://www-seq.wi.mit.edu
         *  Contact: sequence_submissions@genome.wi.mit.edu
         *  ------------------ Project Information
         *  Center project name: L1153
         *  Center clone name: 138_F_3
         *  ------------------ Summary Statistics
         *  Sequencing vector: M13; M77815; 100% of reads
         *  Chemistry: Dye-terminator Big Dye; 100% of reads
         *  Assembly program: Phrap; version 0.960731
         *  Consensus quality: 151183 bases at least Q40
         *  Consensus quality: 160439 bases at least Q30
         *  Consensus quality: 164708 bases at least Q20
         *  Insert size: 167000; agarose-fp
         *  Insert size: 167725; sum-of-contigs
         *  Quality coverage: 4.1 in Q20 bases; agarose-fp
         *  Quality coverage: 4.1 in Q20 bases; sum-of-contigs
         *  ------------------.
         * NOTE: This is a 'working draft' sequence. It currently
         * consists of 28 contigs. The true order of the pieces
         * is not known and their order in this sequence record is
         * arbitrary. Gaps between the contigs are represented as
         * runs of N, but the exact sizes of the gaps are unknown.
         * This record will be updated with the finished sequence
         * as soon as it is available and the accession number will
         * be preserved.
         *        1  1300: contig of 1300 bp in length
         *     1301  1400: gap of    100 bp
         *     1401  2412: contig of 1012 bp in length
         *     2413  2512: gap of    100 bp
         *     2513  4147: contig of 1635 bp in length
         *     4148  4247: gap of    100 bp
         *     4248  5455: contig of 1208 bp in length
         *     5456  5555: gap of    100 bp
         *     5556  6997: contig of 1442 bp in length
         *     6998  7097: gap of    100 bp
         *     7098  8534: contig of 1437 bp in length
         *     8535  8634: gap of    100 bp
         *     8635 10164: contig of 1530 bp in length
         *    10165 10264: gap of    100 bp
         *    10265 10487: contig of 223 bp in length
         *    10488 10587: gap of    100 bp
         *    10588 12420: contig of 1833 bp in length
         *    12421 12520: gap of    100 bp
         *    12521 14843: contig of 2323 bp in length
         *    14844 14943: gap of    100 bp
         *    14944 17265: contig of 2322 bp in length
         *    17266 17365: gap of    100 bp
         *    17366 20071: contig of 2706 bp in length
         *    20072 20171: gap of    100 bp in length
```

```
FEATURES             Location/Qualifiers
     source          1..170425
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1"
                     /clone="RP11-138F3"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1..1300
                     /note="assembly_fragment"
     misc_feature    1401..2412
                     /note="assembly_fragment"
     misc_feature    2513..4147
                     /note="assembly_fragment"
     misc_feature    4248..5455
                     /note="assembly_fragment"
                     vector_side:right"
     misc_feature    10588..12420
                     /note="assembly_fragment"
                     clone_end:SP6
     misc_feature    12521..14843
                     /note="assembly_fragment"
     misc_feature    14944..17265
                     /note="assembly_fragment"
     misc_feature    8635..10164
                     /note="assembly_fragment"
     misc_feature    10265..10487
                     /note="assembly_fragment"
     misc_feature    17366..20071
                     /note="assembly_fragment"
     misc_feature    20172..23531
                     /note="assembly_fragment"
     misc_feature    23632..26137
                     /note="assembly_fragment"
                     clone_end:T7
                     vector_side:left"
     misc_feature    26238..28858
                     /note="assembly_fragment"
```

```
         20172 23531: contig of 3360 bp in length
         23532 23631: gap of    100 bp
         23632 26137: contig of 2506 bp in length
         26138 26237: gap of    100 bp
         26238 29858: contig of 3621 bp in length
         29859 29958: gap of    100 bp
         29959 36225: contig of 6267 bp in length
         36226 36325: gap of    100 bp
         36326 44015: contig of 7690 bp in length
         44016 44115: gap of    100 bp
         44116 49493: contig of 5378 bp in length
         49494 49593: gap of    100 bp
         49594 56796: contig of 7203 bp in length
         56797 56896: gap of    100 bp
         56897 62860: contig of 5964 bp in length
         62861 62960: gap of    100 bp
         62961 72552: contig of 9592 bp in length
         72553 72652: gap of    100 bp
         72653 83310: contig of 10658 bp in length
         83311 83410: gap of    100 bp
         83411 92992: contig of 9582 bp in length
         92993 93092: gap of    100 bp
         93093 103886: contig of 10794 bp in length
        103887 103986: gap of    100 bp
        103987 119418: contig of 15432 bp in length
        119419 119518: gap of    100 bp
        119519 136695: contig of 17177 bp in length
        136696 136795: gap of    100 bp
        136796 153977: contig of 17182 bp in length
        153978 154077: gap of    100 bp
        154078 170425: contig of 16348 bp in length.
```

```
misc_feature    29959..36225
                /note="assembly_fragment"
misc_feature    36326..44015
                /note="assembly_fragment"
misc_feature    44116..49493
                /note="assembly_fragment"
misc_feature    49594..56796
                /note="assembly_fragment"
misc_feature    56897..62860
                /note="assembly_fragment"
misc_feature    62961..72552
                /note="assembly_fragment"
misc_feature    72653..83310
                /note="assembly_fragment"
misc_feature    83411..92992
                /note="assembly_fragment"
misc_feature    93093..103886
                /note="assembly_fragment"
misc_feature    103987..119418
                /note="assembly_fragment"
misc_feature    119519..136695
                /note="assembly_fragment"
misc_feature    136796..153977
                /note="assembly_fragment"
misc_feature    154078..170425
                /note="assembly_fragment"

BASE COUNT    48243 a   34101 c   34753 g   50621 t   2707 others
ORIGIN

Query Match  98.3%;  Score 5181.4;  DB 77;  Length 170425;
Best Local Similarity  99.6%;  Pred. No.0;  Matches 11;  Indels  9;  Gaps  6;
Matches 5257;  Conservative  0;  Mismatches  0;
```

```
Qy        1 atcttctgttcagttta--cctcaggggctacttatgtatgtgaaatgaaatgagataaccaaatgtgaa 58
            |||||||||||||||||   |||||||||  |||||| ||||||||||||||||||||||||||||||||
Db    63685 ATCTTTGTTCAGTTTATATCCTACAGGGCGTATTATGTATATGAAATGAAGATAACCAAATGTGAA 63744

Qy       59 agtcctatataactgtatagcctccattcg-gatgtatgtcttggcagg--atgataaag 115
            |||||||| |||||||||||||||||||||| |||| | ||||||||||||   ||| ||||
Db    63745 AGTCCTATATACTGTATAGCCTCCAATCGTGATGTAGTCTTTGGCAGGTATGATGTATAAAG 63804

Qy      116 aatcaggaaaggaggtatccacgttagccaagtgtcccaggcgtctgtctgtctcttattt 175
            | |||||||||| |||||| ||||||| |||||| ||  ||||||| ||||| ||||||||||
Db    63805 AATCAGGAAAGGAGGTATCCCACGTTAGCCAAGTGTCCAGCCGTCTGTGTCTCGTCTTATTT 63864

Qy      176 agtgacaagtgttgctccctgacagagaagctattcttcaggaaaacatcacaatccaatatggt 235
            |||| |||||||||||||||||||||||||| |||||||||||||||| ||||| ||||||| |||
Db    63865 AGTGACAGATGTTGCTCCTGACAGAGAAGCTATTCTTCAGGAAAACATCACAATCCAATATGGT 63924

Qy      236 aaatccatcaaacaggagctaagaaagaagtggagctgcccaagagaaaat 295
            |||||||||||||||||| ||||||||||||| |||||||||||||||||||
Db    63925 AAATCCATCAAACAGGAGCTAAGAAACAGAGGAATGAGATGGGCACTTGCCCAAGAGAAAAT 63984

Qy      296 gccaggagagcaaataatgatgaaaaataaactttccctttgttctttttcaggaa 355
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    63985 GCCAGGAGAGCAAATAATGATGAAAAATAAACTTTCCCTTTGTTCTTTTTCAGGAA 64044

Qy      356 aaatgatgaggaccaaaatcaatgaatgaaagctccagaaaaaagtgttccaa 415
            |||||||||| ||||||| |||||||||||||||||||||||||||||||||||
Db    64045 AAAATGATGAGGACCAAAATCAATGAATGAAGCTCCAGAAAAAAGTGTTCCAA 64104

Qy      416 attggtaattaagtatttgttccctgggaggagagacctccagcttgatgggaaat 475
            |||||| |||||||| |||||||||||||||||| ||||||||||||| ||||||||
Db    64105 ATTGGTAATTAAGTATTTGTTCCTGGGAGAGACCTCCAGCTTGATGGGAAAT 64164

Qy      476 gggaaaaacgtcaaaagcatcaagctggatctatcattattttaaaac 535
            ||||||||||||||||||||| ||||||||| ||||||| |||||||||
Db    64165 GGGAAAAACGTCAAAAGCATGATCTGATCAGATCCAAAGTGGATTTATTTATTTTAAAAAC 64224

Qy      536 cagatggcatcactctgggggaggcaagttcaggaagtcatgttagcaaggacataaca 595
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    64225 CAGATGGCATCACTCTGGGGAGGCAAGTTCAGGAAGTCATGTTAGCAAAAGGACATAACA 64284
```

```
Qy      596 ataacacgtaaaatcaaaaattccgcaaatgcaggagaggaaaaatgggggactgggaaaagcttc 655
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    64285 ATAACACAGTAAAATCAAAAATTCCGCAAATGCAGGAGAGGAAAAATGGGGACTGGGAAAGCTTTC 64344

Qy      656 ataacagtgattaggtcagttgaccatgttcgcaacacctcccgtctataccaggggaca 715
            |||||||| ||||||||||||||||||||||||||||||||||||| |||||||| |||||
Db    64345 ATAACACGTGATTAGGCAGTTGACCATGTTCGCAACACCTCCCGTCTATACCAGGGAACA 64404

Qy      716 caaaaaattgactgggtcaaggcctaagcctgacttttcaagggaaatgaaaact 775
            ||||||| ||||||||| ||||||||||| |||||||||||||||| |||||||||
Db    64405 CAAAAATTGACTCGGGCCTAAGCCTCGACTTTCAAGGGAAATATGAAAAACTGAGACGAAAA 64464

Qy      776 caaaagacatgttaaaaggcaaccagaacatttgttgagccttcaaagtcagtgccct 835
            |||||||||||| |||||||||||||||||| |||||||||||||||||||||||||||
Db    64465 CAAAAGACATGTTAAAAGGCAACCAGAACATTGTTGAGCCTTCAAAGTCAGTGCCCT 64524

Qy      836 cagcaggggacctgagcgcattgccttaggaaggccagtttcctttaaggaatcttaaga 895
            |||||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||
Db    64525 CAGCAGGGACCCTGAGCGCATTTGCCTTAGGAAGGCCAGTTTTCTTTAAGGAATCTTAAGA 64584

Qy      896 aacttctgaaaatcatgaattttaaccaattttaagtatataaaacaaatatgcgagtcata 955
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    64585 AACTTCTTGAAAGATCATGAATTTTAACCATTTTAAGTATAAAACAAATATGCGAGTCATA 64644

Qy      956 atcagtttagacatgggtcccaattttatataaagtcagcacagcaaggacataacgtgtccca 1015
            ||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    64645 ATCAGTTTAGACATGGGTCCCAATTTTATATAAAGTCAGCACAGGATAACGTGTCCCA 64704

Qy     1016 gctccggatagtcagaaatcattagaaatcactgttgtcccatcctaacttttcagaa 1075
            |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    64705 GCTCCGGATAGGTCAGAAATCATTAGAAATCACTGTTGTCCCATCCTAACTTTTCAGAA 64764

Qy     1076 tgatctcatagcctcacacacacgcccatgtgttctgacctacaaccactctacaa 1135
            ||||| ||||||||||||||||| |||||||||||||||||||| |||||||||||
Db    64765 TGATCTGTCATAGCCCTCACACACGCCCGATGTGTTCTGACCTACAAATGCCACCTACAA 64824

Qy     1136 cccaagtgcctcaaccattgttaacgtgtcatctcagtaggtccccatccgatacaacg 1195
            ||||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||
Db    64825 CCCAAGTGCCTCAACCATTGTTAACGTGTCATCTCAGTAGGTCCCCATCTACAAATGCACC 64884

Qy     1196 tcccctgtgcagacccatccccgctccccacaggaagtctcccacctagactctgcatctt 1255
            ||||||||||||||| ||||||||||||||||||||||||||||||| ||||||||||||
Db    64885 TCCCCTGTGCAGACCCATCCCCGCTCCCCACAGGAAGTCTCCCACTAGACTCTGCATGC 64944

Qy     1256 gatgttacagccagaagctccgtcctgggtgagggagtctgtgtcttgtcttcatcac 1315
            |||||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db    64945 GATGTTACAGCCAGAAGCTCCGTCCTGGGTGAGGGAGTCTGTGTCTTACACCTACCTGAGC 65004

Qy     1316 tctacacctgagctcactgtcaacctctgcctcccaagttcaagcaattctctgttcca 1375
            |||||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db    65005 TCTACACCTGAGCTCACTGTCAACCTCTGCCTCCCAAGTTCAAGCAATTCTCTGTTCTCAG 65064

Qy     1376 cctcccggtagctgggactacaggcgcacgccacctaattttgtttgtttttagtagag 1435
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65065 CCTCCCGGTAGCTGGGACTACAGGCGCGCCCACCTAATTTTGTTTGTATTTGTTAGTAGAG 65124

Qy     1436 atggggtttcaccatattagccgggctggtcttgaactcctggacctcaggtgatccaccc 1495
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65125 ATGGGGTTTCACCATATTAGCCGGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCACC 65184

Qy     1496 acctcagcgcctcaaagtgctgggattacagacatgagtccaggcgccgcccggggtc 1555
            |||||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||
Db    65185 ACCTCAGCCTCCTCAAAGTGCTGGGATTACAGGTCACCGCGCCCACCGGGGTC 65244

Qy     1556 agtgtttaataggaataacttgaatggttctaaaacaaccaacaagggaaacagacaaaagc 1615
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65245 AGTGTTTAATAGGAATAACTTGAATGGTTCTAAAACCAACAACAGACAAAAAGC 65304

Qy     1616 tgtgataaattcagggggattcttgggattggggaatggtgcctccatgagctgcctccagtcc 1675
            ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    65305 TGTGATAAATTTCAGGGATTCTTGGGATGGGGAATGGTGCCATGAGCTGCCTCCCTAGTCC 65364
```

```
Qy     1676 cagaccaactggtcctcaatcactcctcatccatcctcatttcaggcgctaagttaccatt 1735
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65365 CAGACCAACTGGTCCTCAATCACTCCTCATCCATCCTCATTTCAGGCTAAGTTACCATT 65424

Qy     1736 ttattcaccatgctttttgtggtaagcctccacatcgttactgaaataagagtatacataa 1795
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65425 TTATTCACCATGCTTTTTGTGGTAAGCCTCCACATCGTTACTGAAATAAGAGTATACATAA 65484

Qy     1796 actagttccatttttggggcactctgtgtgtgtatgagggagggggaggggcataccccag 1855
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65485 ACTAGTTCCATTTTGGGGCCATCTGTGTGTGTGTATGAGGGAGGGGCATACCCCAGAGA 65544

Qy     1856 ctccttgaagcccccggcgggggttcctctcctccctcgtgtgtgtataggggaggggc 1915
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65545 CTCCTTGAAGCCCCCGGCAGAGGTTTCCTCTCGTGTGTGTATAGGGGAGGGGC 65604

Qy     1916 ggtcctgggtgtcctgagcaagccaatctgtgtgtgtgagggagatggggag 1975
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65605 GGTCCTGGGTGTCCTGAGCAACCTGCCACTGTGTTTTGTTATCACTCT 65664

Qy     1976 ctaggggaacctgttgcttctctgtgtgactcgttcattcatccaaggcatcattga 2035
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65665 CTAGGGACCTGTTGCTTTCTATTTCATCGTGTGAGTCTCATCCAGGCATTCATTGA 65724

Qy     2036 caattattgagtacttatatctgcagacacccagaagacaaaggtgtgagaaaagcagtc 2095
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65725 CAATTTATTGAGTACTTATATCTGCAGACACCAGAAGACAAAAAGGTGTGAGCAAAAGCAGTC 65784

Qy     2096 actgccctaccttcgtgtgtgagtgacagttcagtgtttctcatggaagacgtgcaga 2155
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65785 ACTGCCCTACCTTCGTGTGGAGGTGACAGTTTCTCATGGAAGAGCTGCGAGAAAATTAAT 65844

Qy     2156 agcccaacttaaaacccagtgctgaagagaaaagaaataaaccactcttgaagaattg 2215
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65845 AGCCAGCAACTTAAACCCAGTGCTGAAAGAGAAAAGGAAATAAACCACCATCTTGAAGAATTG 65904

Qy     2216 tgcgcagcacccctcaacaacaggccccaccctccctgcctcatcgtgccc 2275
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    65905 TGCGCAGCCATCCCTTAACACAGGCCACTCCCTGCTGCCTCATCGTGCCC 65964

Qy     2276 ggaggcccccccaagtcttccccaagcgctctccatccaatcaagtcaaccctgcgagct 2335
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65965 GGAGGCCCCCAAGCCCAGTGCTTCTCAAGCCTCTCCATCCAGTCACAGGGCCTGCAGCT 66024

Qy     2336 ggcctgccctcgcttccctgaatcgtcctggtcgtggagtgaactccctggctcc 2395
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66025 GGCCTGCCCTCGCTTCCCGTGAATCGTCCTGAGCTGGAGAACTCCTTGGCTCC 66084

Qy     2396 aggctcccagaaggaaatgttgaggagggaaaatgagggacttgggggaggacagt 2455
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66085 AGGCTCCAGAAATGGAGAGGGAAATGTAGTCTAACGGAGAAATGTGGAGGGACAGT 66144

Qy     2456 gtttccctcagaggaaaggggcctccacgtccacgtggggactgca 2515
            ||||||||||||||||||||||||||||||||||||||||||||||
Db    66145 GTTTCCCTCAGAGGGAAAAGGGCCCTCCACGTCTCCAAGCTCTGCAGCT 66204

Qy     2516 gggagtggggacgctcgggctgctgaagggaaggtgaaaagggcaagg 2575
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    66205 GGGAGTGGGGACGCTGGGGCTGCTGAAGGGCTGAAAAGGCAAGG 66264

Qy     2576 ctgaagctgcccagatgttcagtgtgttcacgtggagttttccgtgttgctcctg 2635
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66265 CTGAACCTGCCCAGATGTTCAGTGTGTTCAGTGTTTTCCGTGTTGCTTCCTG 66324

Qy     2636 tgagccttttatctcttctgcttctggagagaagtccattctgaaggtggatgca 2695
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66325 TGAGCCTTTTATCTTTTCTCTGCTTGGGAGTCAGGGGAGTTTATTTCATGAAGGGATGCA 66384

Qy     2696 gtttcataaagtcagctgttaaaatccaggactgtggggttttccttcaacgaaggccc 2755
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66385 GTTTCATAAAGTCAGCTGTTAAAATTCCAGGGTGTGCAATGGGTTTTTCCTTCACGGAAGGCC 66444

Qy     2756 tttattaatgggaatatagaagcgagctcattcctaggccgcttaattcacggaagaa 2815
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66445 TTTATTTAATGGGAATATAGGAAGCGAGCTCATTTCCTAGGCCGGTTAATTCACGGAAGAA 66504

Qy     2816 gtgactgggagtcttcttcttcatgtcttcttccggcaactactagcttatacg 2875
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66505 GTGACTGGGAGTCTTCTTCATGTCTTCTTGGGCAACTACTAGCTTCCTAAGCCGGTTAATTCACGGAAGAA 66564

Qy     2876 gcttatgcaagacggtcgaaaaaccttggaatcaggagactcggtctttcttcttggttctg 2935
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66565 GCTTATGCAAGACGGTCGAAAAACCTTGGAATCAGGAGACTCGGTTTCTTCTTGGTTCTG 66624

Qy     2936 ccattggttgctgttcgtcgaccgtgggcaagtgtctccttccctgggccatagtctcct 2995
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66625 CCATTTGTTGGCTGTTTCTCATGTCTCCATCAGGGCTGGGCCATAGTCTTCT 66684

Qy     2996 ctgcataaaagaccccttgcagctctctgtctgttccttgggccaatt 3055
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    66685 CTGCTATAAAAGACCCCTTGCAGCTCTCTGTGTTCCTTGGGCCAATT 66744

Qy     3056 aggggggttgctgtgtcgagaggggaaaggagagctgagcactgaagcagccccaagtg 3115
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66745 AGGGGGATGTGAGCAGGGAAGGAGAGCTGAGCAGGGGAAGGAGAGCTGAGCCACGGGAGGTG 66804

Qy     3116 gagggggacaggagaggcaggaaggcagaagagagctgagcagctgagcagagcagca 3175
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66805 GAGGGGGATGTGAGAGGCAGGAGGGAGAGGCAGACCTGCTCCATCAGTGCCCAATT 66864

Qy     3176 gactccggggaccggagccacacaatgctctccaggagagctgagcaccaccacaccccaatt 3235
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66865 GACTCCAGGACCCGGAGCCACAATGCTTCAGAGAGCTGAGCCCCACACAACAGCCACATT 66924

Qy     3236 ttcccttcccttaagcacatagacaatgggcattttgccaataaccaaaaagagaatgcagaa 3295
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66925 TTCCTTCCCTTAAGCACATAGACAATGGGCATTTTGCCAATAACCAAAAAGAATGCAGAGACTAA 66984

Qy     3296 ctggtggttagtcttttgcctggcattcaaaaactgggccagaggtgaaaatgccaga 3355
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66985 CTGGTGGTTAGTCTTTTGCTTTGCCCTGGCATTCAAAAACTGGGCCCAGAGGTGAAAATGCCAGA 67044

Qy     3416 cacggagtgactcgcagcggcaggcaggaggaggagagtagtgtattgtatgacga 3475
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67105 CACGGAGTGACCTGCACCGCGCAGGCAGGGAGGAGGAGGATAGTGTATTGATGACGACGG 67164

Qy     3476 gaagagcagattcattcaaggacttggaattgacacacagggattattagtcacgtgat 3535
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67165 GAAAAGACAGATTCATTCAAGGACGTGGGAATTGACACCAGGGATTATAGTCCACGTGAT 67224

Qy     3536 ctggggggttctaggagggcaggggtctatatgtgggggggaaaatcagtcaagggggaaagtcg 3595
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67225 CCTGGGTTCTAGGAGGGCAGGGGCTTATATTCGTGGGGGAAAATCAGGGAAGTCG 67284

Qy     3596 gggagacctgatttcctaactatactatttttccttctggttaatgtgtcaagtc 3655
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67285 GGAGACCTGATTTCTAAATACTATATATTTTTCCTTTACCAGCCTGTAATTCTGAACGCAAGTC 67344

Qy     3656 acaaggtagtaactgaggctgtaagattactttagttacctcctattttc 3715
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67345 ACAAGGTAGTAACTGAGGCTGTAAGACTGTAAGAGATTACTAGTTTCCTATTAGGAACTCTTTTC 67404

Qy     3716 tctgtgtgaggttagcacgacacaaggcatcacccgtttcttccttttacggaacttttc 3775
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67405 TCTGTGGAGTTAGCACACAAGGCACACACAACCCCGTTTCTTTTAACAGGAAGGAACATTCCT 67464

Qy     3776 aacagagtaaagccaaacagaattcaagcctaggctctgactatatgattggtttttga 3835
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    67465 AACAGGTAAAGCCAAACAGATTCAAGCCTAGGTCTTGCTGACTATATGAATTGGTTTTTGA 67524

Qy     3836 aaaatcatttcagcgatgttactactctgattcagaaaatgagactagtacccttggtc 3895
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT   4
HSMYOC1

```
LOCUS       HSMYOC1      2800 bp    DNA             PRI    21-APR-1998
DEFINITION  Homo sapiens myocilin (GLC1A) gene, promoter and exon 1.
ACCESSION   AF049791
VERSION     AF049791.1   GI:3065670
KEYWORDS    .
SEGMENT     1 of 3
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2800)
AUTHORS     Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C.,
            Alward,W.L.M., Sheffield,V.C. and Stone,E.M.
TITLE       Characterization and comparison of the human and mouse GLC1A
            glaucoma genes
JOURNAL     Genome Res. (1998) In press
REFERENCE   2  (bases 1 to 2800)
AUTHORS     Fingert,J.H., Ying,L., Swiderski,R.E., Nystuen,A.M., Arbour,N.C.,
            Alward,W.L.M., Sheffield,V.C. and Stone,E.M.
TITLE       Direct Submission
JOURNAL     Submitted (23-FEB-1998) Ophthalmology, University of Iowa, 200
            Hawkins Drive, Iowa City, IA 52242, USA
FEATURES             Location/Qualifiers
     source          1..2800
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q21-q31"
     promoter        1..1905
     exon            1906..2509
                     /gene="GLC1A"
                     /gene="GLC1A"
                     /number=1
BASE COUNT      781 a    588 c    673 g    758 t
ORIGIN
```

Query Match   34.2%; Score 1804.4; DB 66; Length 2800;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1838; Conservative   0; Mismatches   1; Indels   3; Gaps   3;

```
Qy  3431 agccaggggaggaagaagaagagaggagatagtgtatgaagaagaacagagattcat 3490
         ||||||||||||||| |||||||||||| ||||||||||||| ||||| |||||||||
Db     1 AGCCAGGGGAGGAGGAGGAAG-AAAAGAGAGGGGATAGTGTATGAGCAAGAAAAGACAGATTCAT 59

Qy  3491 tcaagggcagtgggaattgacaccacaggattatagtccacgttgatcctggggtttctaggag 3550
         |||||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||||
Db    60 TCAAGGGCAGTGGGAATTGACCACCAGGGATTATAGTCCACGTTGATCCTGGGGTTCTAGGAG 119
```

```
Qy  3551  gcagggctatattgtggggggaaaaatcagttcaaggggagtcggtcggggaacctgattct  3610
          ||||||||||||| ||||||||||||||||||||||| | |||||   | ||||||||||||
Db   120  GCAGGGCTATATGTTGTGGGGGGAAAAATCAGTTCAAGGAGAGTCGGGAGAACCTGATTCT   179

Qy  3611  aatactatattttccttacacaagctgagtaactctgagcaagtcacaaggtagtaactg  3670
          |||||||||||||  |||| ||| ||||||||| |||||||||||||||||||||||||||
Db   180  AATACTATATTTTTCCTTTACAAGCTGAGTAATTCTGAGCAAGTCACAAGGTAGTAACTG   239

Qy  3671  aggctgtaggtaactacttagttctctcctattctctattaggaaactcttctgagttagca  3730
          |||||||| | ||||||||||||||    ||||   |||||||||||||||| ||||||||||
Db   240  AGGCTGTAAGATTACTTAGTTCTCCTTATTAGGAAACTCTTTTCTCTGTTGGAGTTAGCA   299

Qy  3731  gcacaagggcaatcccgttcttttctttaacaggagtaaatgtagaagtgagttagcca  3790
          ||||||||||||||||||| ||||| | |||| | | | ||| ||||| |||||||| ||
Db   300  GCACAAGGGCAATCCCGTTTCTTTTTACAGGAGGAAGTGAGTAAAGCCAAA           359

Qy  3791  cagattcaagcctaggtcttgctcagtcatatatgtaggagaagaaaatcattcagcg  3850
          |||||||||||||||||||||||||     |||  |||||||||||||||||||||||||
Db   360  CAGATTCAAGCCTAGGTCTGCTGCTGTTTCTTTTGGATTGTAGGAGAAGAAAATCATTCAGCG  419

Qy  3851  atgttaactactgattcagaaaatgagactgagctgtaaacaaaca  3910
          ||| |||||| |||||||||| | ||||||||||| |||||||||||
Db   420  ATGTTTACTATTCTGATTCAGAAAAATGAGAGTAGTACCTTTCAGCTGTAAACAAACA   479

Qy  3911  cccggttgtaaatgtctcaggtccttaagctttaacacataaagttgctcaaag  3969
          ||||||||||||| |||||||||||| |||||||||||||| |||||||||||||
Db   480  CCCATTTGTAAATGTCTCAGGTTAACTGACCTTAACACATAAATTAGAAATGAA   539

Qy  3970  tctttagagcaaactgtgttctgtttctccacatctggaggtgagtctgcagtttggaa  4029
          | | || |||||||||||||| |||| | |||||||||| ||||||| ||||||||||||
Db   540  TCTTTAGAGCAAAACTGTGTTTCTCCAC-TCTGGAGGTCTGCCAGGTTTGGAA        598

Qy  4030  atattaacttcacaagtattgacactgtgttggtattataacacataaagttgctcaaag  4089
          ||||| ||||||| |||||||||||||||| ||||||||||| |||||||||||||||||
Db   599  ATATTAACTTCACAAGTATTGACACTGTGTTAACACATAAAGTTGCTCAAAG          658

Qy  4090  gcaatcaattctattgtgctaagttactctgcgttttggtatatatttattggct  4149
          |||||||||||||||| ||||||||||||| |||||| |||||| |||||||||||
Db   659  GCAATCATTTATTTCAAGTTACTTCTGACACAGTTTTGGTATATTATTGGCT       718

Qy  4150  attgccattgctttttgtcttttctcctttgggctttattaatgtaaagcaggagttatta  4209
          |||||||||||||||||||| ||||||| |||| |||||||| || ||||||||||||||||
Db   719  ATTGCCATTGCTTTTTGTTTTCTCTTTGGGTTTAATGTAAAGCAGGGATTATTA         778

Qy  4210  acctacagtccagaagcctgtgaattgaatgagaagaaaaattacatttttgttttttac  4269
          |||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db   779  ACCTACAGTCCAGAAAGCCTGTGAATTGAATGAGGAAAAAATTACATTTTTGTTTTTAC     838

Qy  4270  cacttctaactaaatttaacatttattcattcgaatacgtacaaactcaaagtg  4329
          |||||||||||||||||| |||||||||| ||||| ||||||||||||||||||
Db   839  CACCTTCTAACTAAATTTAACATTTATTCCATTGCCAATAACCAATAAACTACAAAGTG     898

Qy  4330  gtaataacagtacctgtgatttgtcattaccaataagaatatcacagacattttatactat  4389
          |||||||||||||||||||||||||| |||| ||||||||| |||||||||||| |||||||
Db   899  GTAATAACAGTACCTGTGATTTTGTCATTACCAATAGAAATCACAGACATTTTATACTAT   958

Qy  4390  attacagttggttgcagatacgttgtaagtgaaatatattatactcaaaactactttgaat  4449
          ||||||||||||||||||  |||||||||| |||||||||||| ||||||||||||||||||
Db   959  ATTACAGTTGGTTGCAGATACGTTGTAAGTGAAATATTATACTCAAAACTACTTTGAAAT  1018

Qy  4450  tagacctccctgctggatcctgttttaacatattaataaacatgtttaaattttgata  4509
          |||||||||||||||| |||||| ||||| |||||| |||||||||||| |||||||||
Db  1019  TAGACCTCCTGGATCTGTTTTTAACATATTAATAAAACATGTTTAAATTTTGATA       1078

Qy  4510  tttcgataatcatatttcattcattgttccttgtaatctatatttttatatatattg  4569
          ||||||||||||||||||||| | || ||||| |||||||||| ||| |||||||||
Db  1079  TTTTGATAATCATATTTCATTCATTGTTCCTTGTAATCTATATTTTATAAATTTCATA    1138

Qy  4570  aaaacatcttctgagaagaggttccccagaattcaccaaatgagtttggcatgcacac  4629
          |||||||||||| ||||||||||| | ||| ||||||||||| ||||||| |||||||
Db  1139  AAAACATCTTTCTGAGAAGAGTTCCCCAGATTCACCAAATGAGGTTCTTGGCATGCACAC   1198

Qy  4630  acacgcggtaagaactgattcacattggtgcctgagatgcaagact
          ||||||||||||||||||||||||||||||||||||||||||||||

Db  1199  ACACGAGTAAGAAACTGATTTTAGAGGGCTAACATTGACATTGGTGCCTGAGATGCAAGACT  1258
          ||||| ||||||| ||||||  ||||| ||||||||| ||||||||||||||||||||||||
Qy  4690  gaaattagaaagttctcccaaagtacacaagtgtttttaaagctgaggggggaa          4749

Db  1259  GAAATTAGAAAGTTCTCCCAAAGATACACAGTTGTTTTTAAAGCTAGGGGGGGAA        1318
          ||||||||||||||||||||||| |||||| |||||||||||||| |||||||||
Qy  4750  atctgccgcttctataggaatgctctccctgggacctggaggttgctgtgtcctgttgttc  4809

Db  1319  ATCTGCCGCTTCTATAGGAATGCTCTCCCTGGGACCTGGAGGTGCTGTCCTTGTGTTC     1378
          ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||| ||||
Qy  4810  tggctggctgttattttttctctctgtcctgctacgtcttaaggaactgttttggatctcca  4869

Db  1379  TGGCTGGCTGTTATTTTTTCTCTCTGCTCCTACGTCTTAAAGGGACTTGTTCTTGGATCTCCA  1438
          ||||||||||||||||||||||||||| |||||||||||| |||| |||| ||||||||||||
Qy  4870  gttccctggccatggtgtgctggcacggttctcaatggtttcaagttgcaggttttggatca  4929

Db  1439  GTTCCCTGGCCATGGTGTCGCTGGCACAGTTCTCAATGAGTTTGCAGAGTGGAA         1498
          |||||||||||||||||| ||||||||| |||||||||   |||||||  |||||
Qy  4930  atataaactagaaaatatatccttgttgaaatcagcagcaccaccagtagtcctgtgttat  4989

Db  1499  ATATAAACTAGAAAATATATCCTTGTTGAAATCAGCACACCAGTGTCCTGTGTAAGTGT    1558
          ||||||||||||||||||||||||||||||||||||| ||||| ||||||| | |||| |
Qy  4990  gttgtacgtgtgtgtgttgtgtgtgtgtgtgtaaaaccagttgggagatataggaacta  5049

Db  1559  GTGTACGTGTGTGTGTTGTGTGTGTGTGTAAAAACCAGTTGTGGGAGAATATAGGAACTA   1618
          |||||||||||||||||||||||||||||    ||||||||  |||| |||||||||||||
Qy  5050  tcattgggggtatggggtcccataaatttgggaattctttttaaaacaaaactccaaacggc  5109

Db  1619  TCATTGGGGTATGGGGTGCGCATAAAATTGGGATGTTCTTTTAAAAGAAAACTCCAAACAGAC  1678
          |||||||||||||||| | |||||| |||||||| ||||||||||| ||| |||||||| | |
Qy  5110  tcctggaaggttatttctctaagaatcttgctctggcagcgtgaggcagcaaccccctgtgcac  5169

Db  1679  TTCTGGAAGGTTATTTTTCTAAGAATCTTGGCTGGGCAGCCTGTGAGAGGCAACCCCCGTGTGCAC  1738
          | |||||||||||||||| |||||||||||| | ||||||  ||||| |||||| ||||| |||||
Qy  5170  agccccccagcctcacgtggccacccctctgtccccagctggggctggcctccccag         5229

Db  1739  AGCCCCCACCCAGCCTCACGTGCGCCACCCTCCCCATGAAAGGGCTGGCTCCCCAG        1798
          |||||| || |||||||||||  ||||||| |||||     |||| ||| ||||||
Qy  5230  tatatataaacctctctcggaggtccggccatgagccacacaagg               5271

Db  1799  TATATATAAAACCTCTCTCGGAGCCTCCGGCGGCGATGACCCAGCAAGG              1840
          |||||||||| ||||||||||||| |||| |   ||||| ||||| ||||
```

RESULT 5

LOCUS   HSGLC1A1   1086 bp   DNA   PRI   30-JAN-1998
DEFINITION Homo sapiens GLC1A (trabecular meshwork induced glucocorticoid
            response) gene, exon I, joined CDS.
ACCESSION   Z97171
VERSION   Z97171.1 GI:2425156
KEYWORDS   GLC1A.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1086)
  AUTHORS   Stone,E.M., Fingert,J.H., Alward,W.L., Nguyen,T.D., Polansky,J.R.,
            Sunden,S.L., Nishimura,D., Clark,A.F., Nystuen,A., Nichols,B.E.,
            Ritch,R., Kalenak,J.W., Craven,E.R. and Sheffield,V.C.
  TITLE   Identification of a gene that causes primary open angle glaucoma
  JOURNAL   Science 275 (5300), 668-670 (1997)
  MEDLINE   97158493
REFERENCE   2 (bases 1 to 1086)
  AUTHORS   Adam,M.F., Belmouden,A., Binisti,P., Brezin,A.P., Valtot,F.,
            Bechetoille,A., Dascotte,J.C., Copin,B., Gomez,L., Chaventre,A.,
            Bach,J.F. and Garchon,H.J.
  TITLE   Recurrent mutations in a single exon encoding the evolutionarily
            conserved olfactomedin-homology domain of TIGR in familial
            open-angle glaucoma
  JOURNAL   Hum. Mol. Genet. 6 (12), 2091-2097 (1997)
  MEDLINE   97472461
REFERENCE   3 (bases 1 to 1086)

```
AUTHORS     Garchon,H.J.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUN-1997) Garchon H.-J; Hopital Necker, INSERM U25,
            161 rue de Sevres 75743 Paris cedex 15 FRANCE
FEATURES             Location/Qualifiers
     source          1..1086
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="q23-q25"
                     /tissue_type="white blood cells"
     TATA_signal     355..363
     exon            385..1065
                     /gene="GLC1A"
                     /label=ex1
                     /number=1
     mRNA            join(385..1065,297177.1:21..146,297174.1:21..1475)
                     /gene="GLC1A"
                     /label=RNAGLA1A
     gene            join(385..1065,297177.1:21..146,297174.1:21..1475)
                     /gene="GLC1A"
     misc_feature    462..1065
                     /gene="GLC1A"
                     /note="first coding exon"
     CDS             join(462..1065,297177.1:21..146,297174.1:21..805)
                     /gene="GLC1A"
                     /label-start
                     /codon_start=1
                     /label=CDSGla1A
                     /protein_id="CAB09899.1"
                     /db_xref="GI:2425157"
                     /db_xref="SWISS-PROT:Q99972"
                     /translation="MRFFCARCCSFGPEMPAVQLLLACLVWDVGARTAOLRKANDQS
                     GRCQYTSVASPNESSSCPEQSQAMSVIHNLQRDSSTQRDLEATKARLSSLSLLHQL
                     TLDOAARPOETOGLQRELGTLRREROLETOTRELETAYSNLLRDKSVLEEKKRLR
                     OENENLARRLESSOEVARLRRGCOPQTRDTARAVPPGSREVSTWNLDTLAFOELKSE
                     LTEVPASRILKESPSGIVRSGEGDTGGELVWWGEPLLRTAETIITGKYGVWMRDPKP
                     TYPYTOETTWRIDTVGTDVROVFEYDILSOFMOGYPSKVHILPRPLESTGAVVYSGSL
                     YFOGAESRTVIRVELNTETVKAEKEIPGAGYHGOFPYSWGGYTDIDLAVDEAGLWVIY
                     STDEAKGAIVLSKLNPENLELEQTWETNIRKQSVANAFIICGLYTVSSYTSADATVN
                     FAVDTGISKLTLIPFKNRYKYSSMIDYNPLEKKLFAWDNLNMVTYDIKLSKM"
BASE COUNT      277 a    282 c    316 g    211 t
ORIGIN
```

```
Qy      4876  agcatagtagctggcacagtgcaggttccaatgcgagtgaatggaaataa 4935
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1  AGCATAGTGCCTGCCACAGTGCAGGTTCTCAATGAGTTTGCAGAGTGAATGGAAATATAA 60

Qy      4936  actagaaatatatcctcttgttgaatcagcacacagtagtcctggtcgtgtcgtgttac 4995
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61  ACTAGAAATATATCCTTGTTGAAATCAGCACACACCAGTAGTCCTGGTGTAAGTGTGTGTAC 120

Qy      4996  gtgtgtgtgtgtgtgtgtgtgtaaaaacaaggtgggaagtaataggaaactattatg 5055
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121  GTGTGTGTGTGTGTGTGTGTCTGCAGCGTGAAGGCAACCCCCTGTGCACAGCCC 180

Qy      5056  gggtgatggggtgcataaaatgggatgttcctttcttttaaaagagaacgacttctcgg 5115
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181  GGGTATGGGTGCATAAAATGGGATGTTCTTTTTTAAAAGAAACTCCAAACAGACTTCTCGG 240

Qy      5116  aaggtatttcttaagaatcttgtcggcagcgtgaaggcaaccccctgtgtcacagcccc 5175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241  AAGGTTATTTCTTAAGAATCTTGCTGGCGAGCCGTGAAGGCAACCCCCTGTGTCACAGCCCC 300

Qy      5176  accaggcctcaccgtcttcccccaagtata 5235
              |||||||||||||||||||||||||||
Db       301  ACCAGGCCTCACGTGGCCAACCTCTGTCTTCCCCATGAAGGGCTGGCTCCCAGTATATA 360
```

```
Qy      5236  taaacctctctgggagctcgggcatgagcgagccaagcaagg 5271
              ||||||||||||||||||||||||||||||||||||||||
Db       361  TAAACCTCTCTGGAGCTCGGGCATGAGCGAGCCACCAAGG 396
```

```
Qy      Best Local Similarity    99.7%;
        Matches   395; Conservative   0; Mismatches   1; Indels   0; Gaps   0;
```

```
Qy      Query Match    7.5%;    Score 394.4;    DB 66;    Length 1086;
Best Match                     Score 394.4;    Pred. No. 9.1e-75;
```

RESULT 6
```
Qy      5236  taaacctctctgggagctcgggcatgagcgagccaagcaagg 5271
              |||||||||||||||||||||||||||||||||||||||
Db       361  TAAACCTCTCTGGAGCTCGGGCATGAGCGAGCCACCAAGG 396
```

```
LOCUS       AB006686S1    1228 bp    DNA                PRI    14-APR-2000
DEFINITION  Homo sapiens gene for myocilin, exon 1, complete sequence.
ACCESSION   AB006686
VERSION     AB006686.1    GI:2828342
KEYWORDS    myocilin.
SEGMENT     1 of 3
SOURCE      Homo sapiens pre-pro-B cell,cell_line:FLEB14-14 DNA, clone_lib:BAC
            clone:KB780C12.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Escribano,J., Ortego,J. and Coca-Prados,M.
  TITLE     Isolation and characterization of cell-specific cDNA clones from a
            subtractive library of the ocular ciliary body of a single normal
            human donor: transcription and synthesis of plasma proteins
  JOURNAL   J. Biochem. 118 (5), 921-931 (1995)
   MEDLINE   96318503
REFERENCE   2  (sites)
  AUTHORS   Stone,E.M., Fingert,J.H., Alward,W.L., Nguyen,T.D., Polansky,J.R.,
            Sunden,S.L., Nishimura,D., Clark,A.F., Nystuen,A., Nichols,B.E.,
            Ritch,R., Kalenak,J.W., Craven,E.R. and Sheffield,V.C.
  TITLE     Identification of a gene that causes primary open angle glaucoma
  JOURNAL   Science 275 (5300), 668-670 (1997)
   MEDLINE   97158493
REFERENCE   3  (sites)
  AUTHORS   Kubota,R., Noda,S., Wang,Y., Minoshima,S., Asakawa,S., Kudoh,J.,
            Mashima,Y., Oguchi,Y. and Shimizu,N.
  TITLE     A novel myosin-like protein (myocilin) expressed in the connecting
            cilium of the photoreceptor: molecular cloning, tissue expression,
            and chromosomal mapping
  JOURNAL   Genomics 41 (3), 360-369 (1997)
   MEDLINE   97312692
REFERENCE   4  (sites)
  AUTHORS   Kubota,R., Mashima,Y., Asakawa,S., Minoshima,S.,
            Hejtmancik,J.F., Oguchi,Y. and Shimizu,N.
  TITLE     Genomic organization of the human myocilin gene (MYOC) responsible
            for primary open angle glaucoma (GLC1A)
  JOURNAL   Biochem. Biophys. Res. Commun. 242 (2), 396-400 (1998)
   MEDLINE   98113364
REFERENCE   5  (bases 1 to 1228)
  AUTHORS   Shimizu,N. and Kudoh,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-1997) to the DDBJ/EMBL/GenBank databases.
            Nobuyoshi Shimizu, Keio University School of Medicine, Department
            of Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582,
            Japan (E-mail:shimizu@dmb.med.keio.ac.jp,
            Tel:81-3-3331-2370(ex.2720), Fax:81-3-3351-2370)
FEATURES             Location/Qualifiers
     source          1..1228
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="FLEB14-14"
                     /cell_type="pre-pro-B cell"
                     /chromosome="1"
                     /clone="KB780C12"
                     /clone_lib="BAC"
                     /map="1q23-q24"
     misc_feature    216..221
                     /note="glucocorticoid responsive element consensus
                     sequence"
     TATA_signal     371..375
     exon            371..1077
                     /number=1
                     /product="myocilin"
BASE COUNT      312 a    325 c    350 g    241 t
```

ORIGIN

```
Query Match
Best Local Similarity  99.3%;  Score 394.4;  DB 9;  Length 1228;
Matches 407;  Conservative 0;  Mismatches 1;  Indels 2;  Gaps 1;  Pred. No. 9.1e-75;

Qy  4862 gatctccagttcctagcatagtgcctggcaacgtcaggttccaagttcagagt       4921
         ||||||||||||||||||||||||||||||  |||||||| ||| ||||||||||||
Db     1 GATCTCCAGTTCCTAGCATAGTGCCTGGCACACGTGCAGGTTCTCAATGAGTTTGCAGAGT   60

Qy  4922 gaatggaaatataaactagaaatatatatccttgttgaatcagcagccctggt       4981
         |||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db    61 GAATGGAAATATAAACTAGAAATATATATCCTTGTTGAATCAGCAGCCAGTAGTCCCTGGT  120

Qy  4982 gtaagtgtgtgacgtgtgtgtgtgtgtgtaaaaccaagtggagatat       5041
         |||||||||||   ||||||||||||||||||||||||| ||||||||||
Db   121 GTAAGTGTGTGTAC--GTGTGTGTGTGTGTGTGTAAAACCAGGTGGAGATAT         178

Qy  5042 aggaactattattgggtgatgggtgtgtcataaaatggagatgttctttttaaaaagaaactcc  5101
         ||||||||||||||||||||||||| |||||||||| |||||||||||||||||| |||||||||
Db   179 AGGAACTATTATTGGGTGATGGGTGCATAAAATGGGAGATGTTCTTTTTAAAAAAGAAACTCC  238

Qy  5102 aaacagacttctggaaggttatttttctaagaaatcttgtcggccgtgaaggcaaccccc  5161
         |||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db   239 AAACAGACTTCTGGAAGGTTATTTTCTAAGAAATCTTGCTCGGCCGTGAAGGCAACCCCC  298

Qy  5162 ctgtgtgcacagcccccacccccagcctcacgtggccaacctctgtcttccccccatgaagggctgg  5221
         ||||| ||||||||| || |||||||||||||||||||| ||||||||||| ||||||||||||||| |
Db   299 CTGTGCACAGCCCCACCCACCCCAGCCTCACGTGGCCACCTCTGTCTTCTTCCCCCATGAAGGGCTCG  358

Qy  5222 ctccccagtatatataaccctctcggggactcggcgcatgaagccagcaagg  5271
         ||||||||||||||||| |||||| || ||||||||||||| ||||||||||
Db   359 CTCCCCAGTATATATAAAACCTCTCTGGAGCTCGGCGCATGAGCCAGCAAGG  408
```

RESULT 7

AC007688/c

LOCUS   AC007688   161577 bp   DNA   PRI   01-SEP-1999
DEFINITION Homo sapiens 12p12-27.2-31.7 BAC RPCI11-392P7 (Roswell Park Cancer
            Institute Human BAC Library) complete sequence.
ACCESSION   AC007688
VERSION   AC007688.15 GI:5815499
KEYWORDS   HTG.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 161577)
  AUTHORS   Muzny,D., Arenson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,
            Dugan,S., Durbin,J., Forcum,J., Garcia,C., Gorrell,J.H.,
            Gorrell,L.L., Hernandez,J., Issar,A., Jackson,L., Kneitz,S.,
            Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichterge,O., Liu,W.,
            Logan,O., Lu,J., Marondel,I., Martinez,C., Merscher,S., Miller,A.,
            Montgomery,K., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,
            Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,
            Vo,Q., Williamson,A., Worley,K.C., Xhang,A.M., Yang,R., Yu,W.,
            Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 161577)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (01-JUN-1999) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 161577)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (01-SEP-1999) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Sep 1, 1999 this sequence version replaced gi:5757565.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-help@bcm.tmc.edu

COMMENT   CLONE LENGTH: This sequence does not necessarily represent the
            entire insert of this clone. Overlapping regions of clones are only
            sequenced and submitted once, so the sequence for the remainder of
            the insert may be found in the record for the adjacent clones.
            Overlapping clones are noted at the beginning and end of the
            Features listing.

            ANNOTATION OF FEATURES:
            STSs are identified using ePCR (Genome Res. 7:541-550) searches
            of a local database that includes entries from dbSTS, GDB, and
            local mapping efforts.
            Repeats are identified using RepeatMasker (A. Smit and P. Green,
            unpublished.) for Human and Mouse sequences.
            Genes and Region of sequence similarity are identified by BLAST
            (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
            EST and cDNA sequences. Genes demonstrate at least two exons
            flanked by consensus splice sites that maintained sequence
            continuity across the splice junctions. Sequences that are not
            identical matches are annotated as similar.

            SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
            standard of double strand coverage with a minimum of 2 clones and 2
            reads with no ambiguities or 2 chemistries with a minimum of 2
            clones and 3 reads with no ambiguities. If the sequence quality for
            a region does not meet this standard, it will be indicated in the
            annotation as Low Coverage.

            QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
            standards - estimated error rate less than 1 per 10,000 bases.
            Reports of lowest quality individual bases and measures of base
            quality are listed below. Description of the metrics can be found
            at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.
            html.

            QUALSTAT-REPORT----------

            -------------- Summary Statistics --------------

            Contig length: 161577
            Phrap values in estimate: 160751
            Average error rate (BCM-Phrap estimate):   0.00163681
            Fraction of Phrap values less than 40 :   0.0376047
            Number of consensus changing edits: 30
            Number of N's in consensus :   0

            -------------- Consensus changing edits --------------

            Position   Original+Context   Edited+Context
            7033   acctccgt(n)ccgcccctt   acctccgt(c)ccgcccct
            4567   aaaaaaagaa(n)ggaaataat   aaaaaaaaaa(a)ggaaataat
            51135   aagcaaagaa(n)aagaaggaa   aagcaaagaaa(a)aagaaggaa
            75582   aaaaaaaaaa(n)aangatgtt   aaaaaaaaaa(a)aaggatgtt
            75585   aaaaaaaana(n)ggtgttcgc   aaaaaaaaaa(a)gggttgttgc
            75667   ctaaagcaga(n)taagatcta   ctaaagcaga(a)taagatcta
            75752   ttttaaatag(g)gcttttgcttt   ttttaaatag(t)gcttttgctt
            84017   gggggaggag(n)aaggaggag   gggggaggag(a)aaggaggag
            85227   tttttgttttt(t)tgtttttt   tttgttttt(t)tgttttttt
            99681   ggagttcaag(n)atgcagtgag   ggagttcaag(a)atgcagtgag
            111307   aatctcttat(n)ccggaattca   aatctcttat(g)ccggaattca
            111374   gtatacatca(n)atcntctttt   gtatacatc(c)atccttttt
            111378   actagtatac(n)tcatttttt   actagtatac(c)atcctttttt
            112621   aaaaaaaaaa(n)ccatcctaga   aaaaaaaaaa(a)ccatcctaga
            135812   attcaccctc(c)tttttttt   attcaccctc(t)ttttttttt
            137207   ttgcaggcac(n)cgccaccacg   ttgcaggcac(a)cgccaccacg
            137218   cgcaccacg(n)tggcctacagt   cgcaccacga(a)tggcctacagt
            145113   caagtgaaa(n)natgtagaat   caagtgaaa(a)gatgtagaat
            145114   caagtgaaan(n)atgtagaata   caagtgaaa(a)gatgtagaat
            145232   ggcaccggtg(n)ctcacnccag   ggcaccggtg(g)ctcacgccag
            145233   ggcaccggtt(n)ctcacgccag   ggcaccggtt(g)ctcacgccag
            145238   ggtgntcac(n)cagtaatcc   ggtgntcac(c)ccagtaatcc
            145321   gcaacatgt(t)nnncccatc   gcaacatggt(g)aaacccatc
```

```
FEATURES             Location/Qualifiers
source               1..161577
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12p13.3"
                     /clone="RPCI11-392P7"
     repeat_region   complement(3..65)
                     /rpt_family="MER4B"
     repeat_region   161..471
                     /rpt_family="AluSx"
     repeat_region   complement(607..702)
                     /rpt_family="MIR"
     repeat_region   703..838
                     /rpt_family="FLAM_C"
     repeat_region   complement(839..916)
                     /rpt_family="MIR"
     repeat_region   complement(1205..1377)
                     /rpt_family="MER104"
     STS             1388..1525
                     /standard_name="WIAF-759-STS"
                     /db_xref="dbSTS:65358"
     repeat_region   1683..1971
                     /rpt_family="AluJo"
     repeat_region   1974..2261
                     /rpt_family="AluJb"
     repeat_region   2486..2593
                     /rpt_family="MIR"
     misc_feature    2610..3214
                     /note="Region: Unigene cluster similar to AA056332 and
     STS             2712..2823
                     /standard_name="SHGC-44583"
                     /db_xref="dbSTS:48451"
     repeat_region   3185..3491
                     /rpt_family="AluSq"
     gene            4509..5919
                     /gene="Human 5-hydroxytryptamine7 receptor isoform b mRNA,
     repeat_region   complement(6397..6537)
                     /rpt_family="L2"
     repeat_region   6967..7158
                     /rpt_family="CCCCG)n"
     repeat_region   7094..7240
                     /rpt_family="(CCG)n"
```

```
     gene            join(7241..7357,17979..18117,20062..20242,23659..23933)
                     /gene="Unigene cluster containing N42239, R09325, and
                     AI380369."
     repeat_region   complement(7770..8250)
                     /rpt_family="MER44B"
     repeat_region   8831..9073
                     /rpt_family="MER102"
     repeat_region   complement(9075..9303)
                     /rpt_family="L2"
     repeat_region   9671..10135
                     /rpt_family="L2"
     repeat_region   10426..10592
                     /rpt_family="LTR33A"
     repeat_region   10625..10689
                     /rpt_family="MIR"
     repeat_region   11094..11322
                     /rpt_family="MIR"
     repeat_region   complement(11985..12403)
                     /rpt_family="L2"
```

```
Query Match        3.6%;    Score 189.8;  DB 10;  Length 161577;
Best Local Similarity 78.5%;  Pred. No. 8e-31;
Matches 227;  Conservative  0;  Mismatches 62;  Indels  0;  Gaps  0;

Qy  1280 agggtgaggtgctgtgtcttaccacctgtatgctctacacctgagctcactgaaac 1339
          ||| || |||||| ||| |||| ||||||||||| | |||||| |||||||| ||||
Db 29160 AGAGTCTCACTCTATTACCCAGGCTGAAGTGCAAGTGGCCAATCTCGGCTCATTGCAAACC 29101

Qy  1340 tctggctcccaggttcaagcaattctccctgtctcagctccgtgagtccgttcaagctc 1399
          |||||||||||||||||||| ||| || ||| ||||| |||||||||||||||||||||||
Db 29100 TCTGGCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAG 29041

Qy  1400 gcgcacgcccggctaattttgtattgtagtagagatggggtttcaccatattaggccg 1459
          ||||||||| ||| ||||||||| || |||||| |||| || | ||||||| ||||||||
Db 29040 GCACACGCCCACCTAATTTTTGTATTTTAGTACAGATGGGGTTTCACCGTATTGGCCAG 28981

Qy  1460 gctggtctttgaactcctgacctcaggtgatccaccaccaccctcctcctaaagtgctggg 1519
          |||||| || || ||| || |||||| | || |||||| ||| | ||||| |||||||||||
Db 28980 GCTGGTCTCAAACTCCTGACCTCAGGTGATCCACCCGCCTCAGCCTCCCTAAAGTGCTGGG 28921

Qy  1520 attacaggcatgagtcaccgcgcccggccaaggggtcagtgtttaataag 1568
          |||||||| ||| |  |||   ||| ||||||| |||||| || |||
Db 28920 ATTACAGGCGTGAGCCACCACGCCCGGCCGGCCCTGAAAGGGTTTTAAAGTAG 28872
```

```
RESULT   8
AC023790/c
LOCUS       AC023790   191123 bp   DNA              HTG        25-JUL-2000
DEFINITION  Homo sapiens chromosome 12 clone RP11-377D9, WORKING DRAFT
            SEQUENCE, 35 unordered pieces.
ACCESSION   AC023790
VERSION     AC023790.16  GI:9438256
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 191123)
  AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barbaria,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
            Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
            Holloway,C., Hosak,H., Jackson,L.R., Jackson,L.E., Jia,Y., Jones,M.,
            Kelly,S., Kondejewski,N., Kong,L., Kovar,C., Leal,B., Li,Z.,
            Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
            Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
            Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
            Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
            Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
            Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R.,
```

```
---------- Distribution of Quality < 40 Bases ----------

 1000|
  900|
  800|
  700|
  600|
# 500|                                          *
bases 400|                              *    *    *
  300|                         *    *    *    *    *
  200|                    *    *    *    *    *    *
  100|          *    *    *    *    *    *    *    *
    0|     *    *    *    *    *    *    *    *    *
          5    10   15   20   25   30   35   40
                    Phrap Value Range
```

```
145522 caacatggtt(n)ncccatct
145523 aacatggttn(n)ncccatct
145324 acatggtgga(a)cccatctct
145363 gttcaccagg(c)ngtgtgggt
145377 gttcaccagg(n)gtgtgggt
145397 tagctacttg(t)tctagtcct
145505 gtggcgtgta(t)tctagtcct
       tagctacttg(g)gaggaggat
       aaaaaaaaaa(a)gaaaaaaaa
```

```
TITLE     Direct Submission
JOURNAL   Unpublished
REFERENCE 2  (bases 1 to 193123)
AUTHORS   Tabor,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wahbah,M.,
          Watlington,S., Weinstock,G., Weinstock,I.R., Williamson,A.,
          Worley,K., Wren,J., Wrensford,G., Yu,W., Zhou,X., Nelson,D. and
          Gibbs,R.
TITLE     Direct Submission
JOURNAL   Submitted (18-FEB-2000) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Jul 25, 2000 this sequence version replaced gi:9255941.
          --------------------- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc-help@bcm.tmc.edu
          --------------------- Project Information
          Center project name: HABM
          Center clone name: RP11-377D9
          --------------------- Summary Statistics
          Assembly program: Phrap; version 0.990329
          Consensus quality: 167304 bases at least Q40
          Consensus quality: 176244 bases at least Q30
          Consensus quality: 181341 bases at least Q20
          Estimated insert size: 180060; sum-of-contigs estimation
          Quality coverage: 0x in Q20 bases; agarose-fp estimation
          Quality coverage: 3.1x in Q20 bases; sum-of-contigs estimation
          NOTE: Estimated insert size may differ from sequence length
          (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
          NOTE: This is a 'working draft' sequence. It currently
          consists of 35 contigs. The true order of the pieces
          is not known and their order in this sequence record is
          arbitrary. Gaps between the contigs are represented as
          runs of N, but the exact sizes of the gaps are unknown.
          This record will be updated with the finished sequence
          as soon as it is available and the accession number will
          be preserved.
          *
          *        1: contig of 21094 bp in length
          *     2105: gap of unknown length
          *     2194: contig of 1 bp in length
          *     2195: gap of unknown length
          *     3930: contig of 12736 bp in length
          *    34030: gap of unknown length
          *    34031: contig of 14288 bp in length
          *    48318: gap of unknown length
          *    48319: contig of 1 bp in length
          *    48320: gap of unknown length
          *    48419: contig of 10381 bp in length
          *    58799: gap of unknown length
          *    58800: contig of 1 bp in length
          *    58899: gap of unknown length
          *    58900: contig of 9924 bp in length
          *    68723: gap of unknown length
          *    68724: contig of 1 bp in length
          *    68823: gap of unknown length
          *    68824: contig of 7250 bp in length
          *    76073: gap of unknown length
          *    76074: contig of 1 bp in length
          *    76173: gap of unknown length
          *    76174: contig of 6695 bp in length
          *    82868: gap of unknown length
          *    82869: contig of 1 bp in length
          *    82968: gap of unknown length
          *    82969: contig of 7078 bp in length
          *    90046: gap of unknown length
          *    90047: contig of 1 bp in length
          *    90146: gap of unknown length
          *    90147: contig of 5995 bp in length
          *    96141: gap of unknown length
          *    96142: contig of 1 bp in length
          *    96241: gap of unknown length
          *    96242: contig of 5632 bp in length
          *   101873: gap of unknown length
          *   101874: contig of 1 bp in length
          *   101973: gap of unknown length
          *   101974: contig of 6029 bp in length
          *   108002: gap of unknown length
          *   108003: contig of 1 bp in length
          *   108102: gap of unknown length
          *   108103: contig of 4801 bp in length
          *   112903: gap of unknown length
          *   112904: contig of 1 bp in length
          *   113003: gap of unknown length
          *   113004: contig of 7206 bp in length
          *   120209: gap of unknown length
          *   120210: contig of 1 bp in length
          *   120309: gap of unknown length
          *   120310: contig of 6044 bp in length
          *   126353: gap of unknown length
          *   126354: contig of 1 bp in length
          *   126453: gap of unknown length
          *   126454: contig of 6435 bp in length
          *   132888: gap of unknown length
          *   132889: contig of 1 bp in length
          *   132989: gap of unknown length
          *   132999: contig of 7149 bp in length
          *   140137: gap of unknown length
          *   140138: contig of 1 bp in length
          *   140237: gap of unknown length
          *   140238: contig of 4394 bp in length
```

```
          *   144631: gap of unknown length
          *   144632: contig of 1 bp in length
          *   144731: gap of unknown length
          *   144732: contig of 5581 bp in length
          *   150312: gap of unknown length
          *   150313: contig of 1 bp in length
          *   150411: gap of unknown length
          *   150412: contig of 3702 bp in length
          *   154114: gap of unknown length
          *   154115: contig of 1 bp in length
          *   154214: gap of unknown length
          *   154215: contig of 4244 bp in length
          *   158458: gap of unknown length
          *   158459: contig of 1 bp in length
          *   158558: gap of unknown length
          *   158559: contig of 2827 bp in length
          *   161385: gap of unknown length
          *   161386: contig of 1 bp in length
          *   161485: gap of unknown length
          *   161486: contig of 2835 bp in length
          *   164320: gap of unknown length
          *   164321: contig of 1 bp in length
          *   164420: gap of unknown length
          *   164421: contig of 3294 bp in length
          *   167714: gap of unknown length
          *   167715: contig of 1 bp in length
          *   167814: gap of unknown length
          *   167815: contig of 3200 bp in length
          *   171014: gap of unknown length
          *   171015: contig of 1 bp in length
          *   171114: gap of unknown length
          *   171115: contig of 2729 bp in length
          *   173843: gap of unknown length
          *   173844: contig of 1 bp in length
          *   173943: gap of unknown length
          *   173944: contig of 2688 bp in length
          *   176631: gap of unknown length
          *   176632: contig of 1 bp in length
          *   176731: gap of unknown length
          *   176732: contig of 1917 bp in length
          *   178648: gap of unknown length
          *   178649: contig of 1 bp in length
          *   178748: gap of unknown length
          *   178749: contig of 1484 bp in length
          *   180232: gap of unknown length
          *   180233: contig of 1 bp in length
          *   180332: gap of unknown length
          *   180333: contig of 2772 bp in length
          *   183104: gap of unknown length
          *   183105: contig of 1 bp in length
          *   183204: gap of unknown length
          *   183205: contig of 1516 bp in length
          *   184720: gap of unknown length
          *   184721: contig of 1 bp in length
          *   184820: gap of unknown length
          *   184821: contig of 1317 bp in length
          *   186137: gap of unknown length
          *   186138: contig of 1 bp in length
          *   186237: gap of unknown length
          *   186238: contig of 1976 bp in length
          *   188213: gap of unknown length
          *   188214: contig of 1 bp in length
          *   188313: gap of unknown length
          *   188314: contig of 1760 bp in length
          *   190073: gap of unknown length
          *   190074: contig of 1 bp in length
          *   190173: gap of unknown length
          *   190174: contig of 1682 bp in length
          *   191855: gap of unknown length
          *   191856: contig of 1 bp in length
          *   191955: gap of unknown length
          *   191956: contig of 1168 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..193123
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-377D9"
                     /chromosome="12"

BASE COUNT    54294 a   41007 c   41291 g   53110 t   3421 others
ORIGIN

Query Match  3.6%;  Score 189.8; DB 77;  Length 193123;
Best Local Similarity  78.5%;  Pred. No. 8e-31;
Matches  227;  Conservative  0;  Mismatches  62;  Indels  0;  Gaps  0;

Qy    1280 agggtgagggtccggtctttacaccctacctgtatgcctccacccctgagcctcactgcaacc 1339
           || ||||||||  ||||||| | || || | ||| |||||||||| |||| || ||||||||
Db  134698 AGAGTCTCACTCATTATTACCCAGGCTGAAGTGCGCGCAATCTCGGCTCATTGCAACC 134639

Qy    1340 tctgcctcccagggttcaaagcaattctcctgtctcagcctcccgcgtagcgtgggactacag 1399
           ||||| |||||||||| | ||| ||||||| ||||||||| ||||| |||||||| ||||||
Db  134638 TCTGCCCTCCCAGGTTCAAGTGATTCCTGCCTTAGCTTCCCGAGTAGCTGGGACTACAG 134579

Qy    1400 gcgcacgcccggctaattttgtattgttagtagagatggggtttcaccattagcccg 1459
           | |||||||||||||||||||||||||||| |||||||||||| ||| |||| | ||
Db  134578 GCACACGCCCAGCTAATTTTGTATTTTTAGTACAGATGGGGTTTGCCGTATTGGCCAG 134519

Qy    1460 gctggtcttgaactctgacctcaggtgatccacccacctcagcctcctaaagtgctggg 1519
           ||||| ||||||||||||||||| | |||||| |||| ||||||| |||||||||||||
Db  134518 GCTGGTCTCAAACTCCTGACCTCAGGTGATCCACCCACCCTCAGCCTCCTAAAGTGCTGGG 134459

Qy    1520 attacagcacgagtcaccgcgcccggccaggggtcagtgttaataag 1568
           ||||||| | ||||||||||||| |||| ||| |||| ||  ||||||
Db  134458 ATTACAGTGTGAGCCACCACGCCCGGCCTGGAAAGGGTTTTAAAGTAG 134410
```

RESULT 9
AC012404

```
LOCUS       AC012404    199722 bp    DNA      HTG      26-MAY-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-39M21 map 15, WORKING DRAFT
            SEQUENCE, 6 unordered pieces.
ACCESSION   AC012404
VERSION     AC012404.4  GI:7107808
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 199722)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-39M21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 199722)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-OCT-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 28, 2000 this sequence version replaced gi:6514012.
            All repeats were identified using RepeatMasker:
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            Smit, A.F.A. & Green, P. (1996-1997)
            ------------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ------------------- Project Information
            Center project name: L2261
            Center clone name: 39_M_21
            ------------------- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 178894 bases at least Q40
            Consensus quality: 191626 bases at least Q30
            Consensus quality: 196426 bases at least Q20
            Insert size: 194000; agarose-fp
            Insert size: 199222; sum-of-contigs
            Quality coverage: 6.2 in Q20 bases; agarose-fp
            Quality coverage: 6.1 in Q20 bases; sum-of-contigs
            --------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            *        1  5729: contig of 5729 bp in length
            *     5730  5829: gap of 100 bp
            *     5830  20896: contig of 15067 bp in length
            *    20897  20996: gap of 100 bp
            *    20997  52603: contig of 31607 bp in length
            *    52604  52703: gap of 100 bp
            *    52704  89348: contig of 36645 bp in length
            *    89349  89448: gap of 100 bp
            *    89449 138760: contig of 49312 bp in length
```

```
FEATURES             Location/Qualifiers
     *    138761 138860: gap of        100 bp
     *    138861 199722: contig of 60862 bp in length.
     source          1..199722
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /clone="RP11-39M21"
                     /clone_lib="RPCI-11 Human Male BAC"
                     /map="15"
     misc_feature    1..5729
                     /note="assembly_fragment"
     misc_feature    5830..20896
                     /note="assembly_fragment"
     misc_feature    20997..52603
                     /note="assembly_fragment"
     misc_feature    52704..89348
                     /note="assembly_fragment"
                     /vector_side:left"
     misc_feature    89449..138760
                     /note="assembly_fragment"
     misc_feature    138861..199722
                     /note="assembly_fragment"
                     /clone_end:SP6
                     /vector_side:right"
BASE COUNT    60442 a   42775 c   40458 g   55547 t      500 others
ORIGIN
```

```
Qy    1290 tctgtgtcttacactactgtatgctctacacctgagctcactgcaacctctgcctcc 1349
           |||||  ||||||| ||| ||| | |||||||||||||||||||||||||||||||||
Db  184619 TCTGTCGCCCAGGCTGGAGTGCAANTGGCGCAAATCTACCTCCCTGCCAACCTCTGCCTCC 184678

Qy    1350 aggttcaagcaattctcctgtctccagcctccggtagctgggactacaggcgcacgccc 1409
           ||||||||||| | |||||||||||||||||||||||||||||||  |||| |||||||
Db  184679 AGGTTCGGGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAAGTGCACACT 184738

Qy    1410 ggctaattctttgtattgttagtagagatgggtttcacatattagacccggcctg 1469
           ||||||| ||||||||| |||||||||||| |||||||  |||  |||| ||||||
Db  184739 GGCTAANTTTTGTATTTTTAGTAGAGANAAGGGTTTCACCATATTGGCCAGGGTGGTCTTG 184798

Qy    1470 aactcctgacctcaggtgatccacccaccctccagcctcggggattacaggca 1529
           ||||||||||||||||||||| |||||||||||||||||| |||||||||||||
Db  184799 AACTCCTGACCTCAGGTGATCCACCCACCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCG 184858

Qy    1530 tgagtcaccgcgcccgcgcaagggtcagtgttaata 1566
           ||||| ||||| |||  ||||||||| ||||||||||
Db  184859 TGAGCCACCATGCCAGGCCCATTCTTACTTTTTACA 184895
```

```
RESULT 10
AC004973/c
LOCUS       AC004973    97037 bp    DNA      PRI     21-DEC-1999
DEFINITION  Homo sapiens PAC clone RP5-113911 from Xq23, complete sequence.
ACCESSION   AC004973
VERSION     AC004973.1  GI:3694660
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 97037)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 97037)
  AUTHORS   Courtney,L., Langston,Y. and Drone,K.
```

Query Match   3.6%; Score 189; DB 71; Length 199722;
Best Local Similarity   80.1%; Pred. No. 1.2e-30;
Matches 222; Conservative   0; Mismatches 55; Indels   0; Gaps   0;

```
TITLE       The sequence of Homo sapiens PAC clone RP5-1139I11
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 97037)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (03-OCT-1998) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   5  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAR-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 97037)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (21-DEC-1999) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 3, 1998 this sequence version replaced gi:3213024.
            -------------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@watson.wustl.edu
            -------------- Summary Statistics
            Center project name: H_DJ1139I01
            --------------
            NOTICE:  This sequence may not represent the entire insert of this
            clone. It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            This sequence was generated from part of bacterial clone contigs of
            human chromosome X, constructed by the chromosome X mapping group
            at the Sanger Centre, Wellcome Trust Genome Campus, Hinxton, UK.
            Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX/

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-5, prepared by
            Pieter de Jong and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu) using the method described by
            Ioannou et al., Nature Genetics 6:84-9 (1994).  The library is from
            one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR:  pCYPAC2
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP4-555N2; the clone sequenced
            to the right is RP3-404F18, 200 bp overlap. Actual start of this
            clone is at base position 1 of RP5-1139I11; actual end is at 13269
            of RP3-404F18.
FEATURES             Location/Qualifiers
     source          1..97037
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
```

```
     repeat_region   /map="Xq23"
                     /clone="RP5-1139I11"
                     /clone_lib="RPCI-5"
     repeat_region   530..801
                     /rpt_family="L2"
     repeat_region   830..1035
                     /rpt_family="Alu"
     repeat_region   2179..2349
                     /rpt_family="MIR"
     repeat_region   3434..3489
                     /rpt_family="MER1_type?"
     repeat_region   3526..3739
                     /rpt_family="MIR"
     repeat_region   4681..4852
                     /rpt_family="MIR"
     repeat_region   5395..5961
                     /rpt_family="MER4-group"
     repeat_region   6293..6393
                     /rpt_family="L1"
     repeat_region   6394..7105
                     /rpt_family="L1"
     repeat_region   7104..7211
                     /rpt_family="L1"
     repeat_region   7342..7520
                     /rpt_family="L1"
     repeat_region   7542..7644
                     /rpt_family="Alu"
     repeat_region   7645..8674
                     /rpt_family="Retroviral"
     repeat_region   8682..9314
                     /rpt_family="Retroviral"
     repeat_region   9315..9368
                     /rpt_family="Retroviral"
     repeat_region   9371..10157
                     /rpt_family="7SLRNA"
     repeat_region   10165..10817
                     /rpt_family="Retroviral"
     repeat_region   10818..10973
                     /rpt_family="Alu"
     repeat_region   10974..11111
                     /rpt_family="L1"
     repeat_region   11238..11612
                     /rpt_family="L1"
     repeat_region   11972..12283
                     /rpt_family="MaLR"
     repeat_region   12570..12854
                     /rpt_family="Alu"
     repeat_region   12859..13433
                     /rpt_family="L1"
     repeat_region   13435..13564
                     /rpt_family="MER1_type"
     repeat_region   13580..13845
                     /rpt_family="Alu"
     repeat_region   13850..14418
                     /rpt_family="MER1_type"
     repeat_region   14119..14411
                     /rpt_family="L1"
     repeat_region   14412..14813
                     /rpt_family="Alu"
     repeat_region   14801..15547
                     /rpt_family="L1"
     repeat_region   16501..16554
                     /rpt_family="L1"
     repeat_region   17412..17533
                     /rpt_family="L2"
     repeat_region   17534..17793
                     /rpt_family="MER1_type"
     repeat_region   17794..17912
                     /rpt_family="Alu"
     repeat_region   17933..18249
                     /rpt_family="MER1_type"
     repeat_region   18293..18358
                     /rpt_family="L1"
     repeat_region   18293..18358
                     /rpt_family="MIR"
```

```
RESULT 11
LOCUS       HUMYWXD703    135038 bp    DNA             PRI       24-DEC-1996
DEFINITION  Homo sapiens ADP/ATP carrier protein (ANT-2) gene, complete cds.
ACCESSION   L78810
```

Query Match   3.6%; Score 187.8; DB 9; Length 97037;
Best Local Similarity 78.7%;
Matches 240; Conservative 0; Mismatches 57; Indels 8; Gaps 1;

Pred. No 2.1e-30;

```
Qy  1280 agggtgaggtccggtcttacacctacctgtatgtctacacctgagctcactgcaacc 1339
         ||||||| ||||||||||||||||||||||| |||||||||||||||||||||||||
Db 54922 AGAGTCTTGCTCTGTCATCCAACCTGGAGTGCACGAGATCTCACCTCACTGCAACC 54863

Qy  1340 tctgcctccaggtcaagacaattctccctgtctcagccctccggtagctgggactacag 1399
         ||||||  ||||||||||||| |||||||||||||||||||||||||
Db 54862 TCTGCCTCCTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCAGTAGCTGGGACTACAG 54803

Qy  1400 gcg--------cacgcccggctaacttttttgtattgttagtagagatggggtttcaccata 1451
         |||       |||||||||||||||||||||||| |||||||||||
Db 54802 GCCCATGCCACCACGCCCAGCTAATTTTTGTATTTTAGTAGAGATGGGGTTTCACCATA 54743

Qy  1452 ttagccccggctcggtcttgaacctcctgacctcaggtgatccacccgccctccagcctcctaa 1511
         ||| |||||||||||||||||||||||||||||||| ||||||
Db 54742 TTGGTCAGGCTGGTCTTGAAACTCCTGACCTCAGATGATCCACCCGCCTCCCAAA 54683

Qy  1512 gtgctgtggattacaggcatgagtcacccgcgcccggccaagggtcagtgcttaataaggaa 1571
         ||||||||||||||||||||| |||||||||| ||||| |||||
Db 54682 GTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTCTTCTAGGTATTTTGAAATGTA 54623

Qy  1572 taact 1576
         ||  |
Db 54622 CAATT 54618
```

```
repeat_region   18348..18476  /rpt_family="L2"
repeat_region   18483..18908  /rpt_family="L2"
repeat_region   18932..19227  /rpt_family="MaLR"
repeat_region   19390..19604  /rpt_family="Alu"
repeat_region   19606..20227  /rpt_family="L2"
repeat_region   20228..20477  /rpt_family="AcHobo"
repeat_region   20478..20721  /rpt_family="Alu"
repeat_region   20738..20828  /rpt_family="AcHobo"
repeat_region   20848..21707  /rpt_family="AcHobo"
repeat_region   21717..21711  /rpt_family="L2"
repeat_region   22163..22378  /rpt_family="L2"
repeat_region   22866..23031  /rpt_family="L2"
repeat_region   23060..23339  /rpt_family="Alu"
repeat_region   23409..24201  /rpt_family="L2"
repeat_region   24203..24554  /rpt_family="AcHobo"
repeat_region   24572..25105  /rpt_family="AcHobo"
repeat_region   25106..25479  /rpt_family="L1"
repeat_region   25480..25768  /rpt_family="Alu"
repeat_region   25769..25985  /rpt_family="L1"
```

```
VERSION     L78810.1  GI:1381111
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135038)
  AUTHORS   Chen,C.N., Su,Y., Baybayan,P., Siruno,A., Nagaraja,R.,
            Mazzarella,R., Schlessinger,D. and Chen,E.
  TITLE     Ordered shotgun sequencing of a 135 kb Xq25 YAC containing ANT2 and
            four possible genes, including three confirmed by EST matches
  JOURNAL   Nucleic Acids Res. 24 (20), 4034-4041 (1996)
  MEDLINE   97078684
COMMENT     Submitted by:
            Ellson Chen,
            Advanced Center for Genetic Technology,
            Applied Biosystems Division of Perkin Elmer Corp., 850 Lincoln
            Center Drive,
            Foster City, CA 94404 USA
            and
            David Schlessinger,
            Department of Molecular Microbiology and Center for Genetics in
            Medicine
            Washington University School of Medicine,
            St. Louis MO 63110 USA
            e-mail: ellson@genseq.apldbio.com and davids@genetics.wustl.edu
            Note: Gene predictions were accomplished with runs of Grail
            versions 1.1 and 1.2, coupled with fasta and blastx comparisons to
            genbank & non-redundant peptide libraries. Repeat analysis was
            accomplished via censor.
            The Rat EST105369 shows significant homology via blastx to this
            sequence
            top     35858   36081
            Strand  Start   End
            top     28802   29076
            Strand  Start   End
            top      9072    9326
            The Graves Disease carrier protein (X66035) shows significant
            homology via blastx to this sequence
            Strand  Start   End
            Start   End
            This gene shows homology via blastx to the EST clone y115a12 Strand
            bottom  125256 125915.
FEATURES             Location/Qualifiers
     source          1..135038
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="X"
     repeat_unit     complement(271..560)
                     /rpt_family="Alu-Sx"
                     /evidence=experimental
     repeat_unit     complement(774..1726)
                     /rpt_family="L1"
                     /evidence=experimental
     repeat_unit     complement(1842..2131)
                     /rpt_family="L1"
                     /evidence=experimental
     repeat_unit     2312..2465
                     /rpt_family="L1PA11"
                     /evidence=experimental
     repeat_unit     complement(2468..2757)
                     /rpt_family="Alu-Sx"
                     /evidence=experimental
     repeat_unit     2774..3059
                     /rpt_family="MLT1C"
                     /evidence=experimental
     repeat_unit     complement(3204..3325)
                     /rpt_family="MIR2"
                     /evidence=experimental
     repeat_unit     3384..3674
                     /rpt_family="Alu-Sq"
                     /evidence=experimental
     repeat_unit     3794..4216
                     /evidence=experimental
```

```
repeat_unit      /rpt_family="L1MA9"
                 /evidence=experimental
                 7502..7792
repeat_unit      /rpt_family="Alu-J"
                 /evidence=experimental
                 7965..8072
repeat_unit      /rpt_family="L1MA2"
                 /evidence=experimental
                 8073..8360
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 8379..8535
repeat_unit      /rpt_family="L1MA5"
                 /evidence=experimental
                 8540..8743
repeat_unit      /rpt_family="L1"
                 /evidence=experimental
                 8744..8869
repeat_unit      /rpt_family="Alu-J"
                 /evidence=experimental
                 complement(9771..10055)
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 10292..10780
repeat_unit      /rpt_family="L1"
                 /evidence=experimental
                 11403..11692
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 12097..12459
repeat_unit      /rpt_family="MST"
                 /evidence=experimental
                 complement(12492..12701)
repeat_unit      /rpt_family="Alu-J or an Alu-S"
                 /evidence=experimental
                 complement(12730..13020)
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 13708..13997
repeat_unit      /rpt_family="Alu-Sq"
                 /evidence=experimental
                 complement(14402..14541)
repeat_unit      /rpt_family="MIR2"
                 /evidence=experimental
                 14546..14836
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 15348..15628
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 complement(15805..16906)
repeat_unit      /rpt_family="LTR12"
                 /evidence=experimental
                 17588..17930
repeat_unit      /rpt_family="MER44"
                 /evidence=experimental
                 complement(20000..20276)
repeat_unit      /rpt_family="Alu-Sb0 or an Alu-Sb1"
                 /evidence=experimental
                 20355..20856
repeat_unit      /rpt_family="L1ME3A"
                 /evidence=experimental
                 20900..20986
repeat_unit      /rpt_family="MSTA"
                 /evidence=experimental
                 20988..21276
repeat_unit      /rpt_family="Alu-Sx"
                 /evidence=experimental
                 21342..21723
repeat_unit      /rpt_family="L1ME3A"
                 /evidence=experimental
                 21841..22129
repeat_unit      /rpt_family="Alu-Sp"

repeat_unit      /evidence=experimental
                 22150..22445
                 /rpt_family="Alu-J"
repeat_unit      /evidence=experimental
                 complement(22542..22789)
                 /rpt_family="THE1B"
repeat_unit      /evidence=experimental
                 complement(23235..23377)
                 /rpt_family="MIR2"
repeat_unit      /evidence=experimental
                 complement(23513..23798)
                 /rpt_family="Alu-Sx"
repeat_unit      /evidence=experimental
                 24296..24578
                 /rpt_family="Alu-Sx"
repeat_unit      /evidence=experimental
                 24808..24944
                 /rpt_family="Alu-Sq"
repeat_unit      /evidence=experimental
                 24945..25234
                 /rpt_family="Alu-Sb2"
repeat_unit      /evidence=experimental
                 25236..25393
                 /rpt_family="Alu-Sc"
repeat_unit      /evidence=experimental
                 complement(26486..26773)
                 /rpt_family="Alu-Sx"
repeat_region    /evidence=experimental
                 26570..26859
                 /rpt_family="CpG Island"
repeat_unit      /evidence=experimental
                 26801..27091
                 /rpt_family="Alu-Sx"
repeat_unit      /evidence=experimental
                 27866..27986
                 /rpt_family="MIR"
repeat_unit      /evidence=experimental
                 complement(28015..28185)
                 /rpt_family="MIR"
repeat_region    /evidence=experimental
                 28550..29215
                 /rpt_family="CpG Island"
repeat_unit      /evidence=experimental
                 31026..31123
                 /rpt_family="MLR"
repeat_unit      /evidence=experimental
                 complement(32698..32984)
                 /rpt_family="Alu-Sc"
repeat_unit      /evidence=experimental
                 33130..33399
                 /rpt_family="MER21"
```

```
Query Match         3.6%;  Score 187.8;  DB 67;  Length 135038;
Best Local Similarity  78.7%;      Pred. No. 2.1e-30;
Matches  240;  Conservative  0;  Mismatches  57;  Indels  8;  Gaps  1;

Qy   1280 agggtgagggtctgtgttcttacacctacctgtatgctcagcctcactgcaacc 1339
          | | |  | |   ||| | | | |  | |   |  ||  | ||||||||||||||||
Db   8349 AGAGTCTTGCTCGTTCATCCAAGCTGGAGTGCACGATCTCAGCTCACTGCAACC 8290

Qy   1340 tctgcctcccccagggttcaaagcaattctcctgtcctcagcctcccgtggg... 1399
          |||||||| |||||||| ||||||||||| || |||||||||||||||| |||||
Db   8289 TCTGCCTCCTGGGTTCAAGCAATTCTCCTGTCTCAGCCTCCCGAGTAGCTGGGACTACAG 8230

Qy   1400 gcg------cacgcccggctaattttttgttattgttagtagagatgggg... 1451
          |||      |||||||||||||||||||||| |||||||||||||||||||||
Db   8229 GCGGCATCCCACACGCCCAGCTAATTTTTGTTATTTTTGTATTTTGTAGAGATGGGG 8170

Qy   1452 ttagccccggctggtcttgaacctccagggtgatccacccacccagcctcctaaa 1511
          |||| ||  |||||||||||||||||||||||||||| ||||||||||||||||||
Db   8169 TTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGATGATCCACCTGCCTCGGCCTCCAAA 8110

Qy   1512 gtgctgggattacaggcatgagtcaccgcgcccggccaagggtcagtgttaataggaa 1571
```

```
Db  8109  GTGCTGGGATTACAGGCGTGAGCCACCGCCGCCCGGCCCTTCTAGGTATTTGGAAGTA  8050
          |||||||||| |||||||||| |||| | ||| | |||||||| ||| | ||| |
Qy  1572  taact  1576
          | | |
Db  8049  CAATT  8045
```

RESULT 12

```
LOCUS       HS821D11              76727 bp    DNA              PRI      12-DEC-1999
DEFINITION  Human DNA sequence from clone RP5-821D11 on chromosome 22q12.3-13.1
            Contains three partial unknown genes, one downstream of a predicted
            CpG island, and the first coding exon of the SREBF2 gene for Sterol
            Regulatory Element Binding Transcription Factor 2 downstream of a
            predicted CpG island, ESTs, STSs, GSSs and genomic marker D22S1157,
            complete sequence.
ACCESSION   AL021453
VERSION     AL021453.1  GI:3413288
KEYWORDS    HTG; D22S1157; SREBF2; Sterol Regulatory Element Binding
            Transcription Factor 2.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76727)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 12, 1998 this sequence version replaced gi:3355590.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP5-821D11 is
            from the library RPCI-5 constructed at the Roswell Park Cancer
            Institute by the group of Pieter de Jong. For further details see
            http://bacpac.med.buffalo.edu/.
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP5-821D11 It may be shorter because we only sequence overlapping
            sections once, or longer because we arrange for a small overlap
            between neighbouring submissions.
            The true left end of clone RP5-821D11 is at 1 in this sequence. The
            true left end of clone CTA-250D10 is at 76628 in this sequence. The
            true right end of clone CTA-109G6 is at 42082 in this sequence. The
            start of this sequence overlaps with sequence Z99716 The end of
            this sequence overlaps with sequence Z83840.
FEATURES             Location/Qualifiers
     source          1..76727
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q12.3-13.1"
                     /clone="RP5-821D11"
                     /clone_lib="RPCI-5"
     repeat_region   350..434
                     /note="L2 repeat: matches 2490..2537 of consensus"
     repeat_region   435..743
                     /note="L2 repeat: matches 2381..2490 of consensus"
     repeat_region   446..448
                     /gene="dJ821D11.1"
                     /note="AluJb repeat: matches 8..312 of consensus"
     variation       744..840
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
                     tgt in this entry
                     /replace="ttt"
                     substitution"
     mRNA            join(1230..1711,13695..>13736)
                     /gene="dJ821D11.1"
                     /product="dJ821D11.1 (PUTATIVE protein)"
                     /note="match: ESTs: Em:AA419437"
                     /evidence=not_experimental
     gene            1230..13736
                     /gene="dJ821D11.1"
     variation       1314..1316
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
                     tct in this entry
                     /replace="tct"
                     substitution"
     CDS             join(1601..1711,13695..>13736)
                     /gene="dJ821D11.1"
                     /note="other possible startcodon at 1583
                     this gene and dJ821D11.2 could be part of one gene
                     match: proteins: Tr:O95505"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="dJ821D11.1 (PUTATIVE protein)"
                     /protein_id="CAA16279.1"
                     /db_xref="GI:4200330"
                     /db_xref="SPTREMBL:O95505"
                     /translation="MLALTLAKADSPRTALLCSAWLLTASFSAQQHKGSLQKDPLLSQ
                     ACVGCLE"
     variation       1760..1762
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
                     gag in this entry
                     /replace="gaa"
                     substitution"
     variation       1901..1903
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
                     tca in this entry
                     /replace="tta"
                     substitution"
     repeat_region   1773..1923
                     /gene="dJ821D11.1"
                     /note="MIR repeat: matches 47..201 of consensus"
     repeat_region   2010..2209
                     /note="L2 repeat: matches 2250..2453 of consensus"
     variation       2302..2304
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
                     gga in this entry
                     /replace="gaa"
                     substitution"
     repeat_region   3042..3152
                     /note="L1MC1 repeat: matches 6231..6332 of consensus"
     repeat_region   3162..3457
                     /note="AluSp repeat: matches 1..297 of consensus"
     repeat_region   3458..3551
                     /note="MIR repeat: matches 74..148 of consensus"
     repeat_region   3552..3848
                     /note="AluJb repeat: matches 1..299 of consensus"
     repeat_region   3849..3887
                     /note="MIR repeat: matches 32..74 of consensus"
     variation       4753..4754
                     /gene="dJ821D11.1"
                     /note="Clone CTA-109G6
```

```
repeat_region   tt in this entry
                deletion"
    4772..4790
    /replace="ttgagt"
repeat_region   4791..5099
    /note="LiM4 repeat: matches 3167..3185 of consensus"
    /note="AluSq repeat: matches 2..312 of consensus"
    4791..4792
    /note="AluSq repeat: matches 2..312 of consensus"
variation   4791..4792
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    att in this entry
    substitution"
    /replace="act"
    5430..5729
variation   /replace="act"
repeat_region   5430..5729
    /note="clone CTA-109G6
    tt in this entry
    deletion"
    5100..5116
    /replace="ttttatttat"
variation   5100..5116
    /gene="LiM4 repeat: matches 3151..3167 of consensus"
    /note="clone CTA-109G6
    5107..5112
repeat_region   5107..5112
    /note="LiM4 repeat: matches 3151..3167 of consensus"
    ttttat in this entry
    insertion"
    5117..5427
    /replace="tt"
repeat_region   5117..5427
    /note="AluYb8 repeat: matches 1..318 of consensus"
    5359..5361
variation   5359..5361
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    att in this entry
    substitution"
    5430..5729
    /replace="act"
repeat_region   5430..5729
    /note="clone CTA-109G6
    /note="AluSp repeat: matches 1..300 of consensus"
    5730..5764
repeat_region   5730..5764
    /note="LiM4 repeat: matches 3116..3151 of consensus"
    5761..5762
variation   5761..5762
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    at in this entry
    deletion"
    /replace="att"
    5765..6075
repeat_region   5765..6075
    /note="AluY repeat: matches 1..311 of consensus"
    5772..5773
variation   5772..5773
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    tt in this entry
    deletion"
    /replace="tat"
    5778..5779
variation   5778..5779
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    tat in this entry
    substitution"
    /replace="tt"
    5780..5782
variation   5780..5782
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    cgg in this entry
    substitution"
    5799..5801
    /replace="ttat"
variation   5799..5801
    /gene="dJ821D11.1"
    /note="clone CTA-109G6
    cgg in this entry
    substitution"
```

```
RESULT 13
AC026395/c
LOCUS       AC026395  152044 bp  DNA      HTG  15-JUN-2000
DEFINITION  Homo sapiens chromosome 10 clone RP11-45D20, WORKING DRAFT
            SEQUENCE, 44 unordered pieces.
ACCESSION   AC026395
VERSION     AC026395.3  GI:8567738
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152044)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152044)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jun 15, 2000 this sequence version replaced gi:7330305.
            -------------- Genome Center
            Center: Genome Therapeutics Corporation
            Center code: GTC
            Web site: http://www.genomecorp.com/
            Contact: gtc-seqcenter@genomecorp.com
            -------------- Project Information
            Center project name: hg344
            -------------- Summary Statistics
            Sequencing vector: N/A
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 990315
            Consensus quality: 126801 bases at least Q40
            Consensus quality: 138138 bases at least Q30
            Consensus quality: 140422 bases at least Q20
            Insert size: 147744; sum-of-contigs
            Quality coverage: 3.2x in Q20 bases; sum-of-contigs
            ----------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 44 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            ----------------------------------
            *                    1007    1006: contig of 1006 bp in length
            *                    1007    1106: gap of unknown length
            *                    2476    2475: contig of 1369 bp in length
            *                    2476    2575: gap of unknown length
```

```
Qy      1340 tctgcctcccaaggttcaagcaattcctcgtgtccaagctcccgcgtagctggagtacag 1399
             ||||||||||||||||||||||||| |||||||||   | ||| | ||| |||||||| ||
Db     57358 TCTGCCTCCCAAGGTTCAAGTAAGTAGATTCCTCCGTCTTCGCCTCCGGGGTTCGGGACTATAG 57299
Qy      1400 gcg---------cacgccgggctaattttgtattgttagtagagatggggttcaccata 1451
             ||| ------ |||||||||||||||||||||||||||||||||||||||||||||
Db     57298 GCCGTGTACCACCATGCCCGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATA 57239
Qy      1452 ttagcccggctggtcttgaactcctgacctcaggtgatccacccacctcagcctcccaaa 1511
             ||-|||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     57238 TTGGCTAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAA 57179
Qy      1512 gtgctgggattacaggcgtgagccaccgcgcccggccaggcaggtcgtttaataaggaa 1571
             |||||||||||||||||||||||||| ||||||||| |||||||||||| || ||||||
Db     57178 GTGCTGGGATTACAGGCGTGAGCCACTGTGCCCGGCCAAAAATCTGGGGTATTTCAAATAA 57119
Qy      1572 taa 1574
             ||
Db     57118 AAA 57116
```

```
Query Match        3.5%;  Score 185.8;  DB 65;  Length 76727;
Best Local Similarity   78.5%;  Pred. No. 5.6e-30;
Matches  238;  Conservative  0;  Mismatches  57;  Indels  8;  Gaps  1;

Qy      1280 agggtggagggtctgtgtctcacctacctgtatgtctcctaccctcactgcaacc 1339
             ||| ||||| | || || | ||  |||  |||| | ||| | |||||| ||||||
Db     57418 AGAGTCTCATTCATTCTGTTGCCTAGGCTGGAGTGCAGTGGCGATGATCTCAGCTTACTGCAACC 57359
```

```
 2576   3927:  contig of 1352 bp in length
 3928   4027:  gap of unknown length
 4028   5333:  contig of 1306 bp in length
 5334   5433:  gap of unknown length
 5434   6564:  contig of 1131 bp in length
 6565   6664:  gap of unknown length
 6665   7870:  contig of 1206 bp in length
 7871   7970:  gap of unknown length
 7971   9384:  contig of 1414 bp in length
 9385   9484:  gap of unknown length
 9485  10791:  contig of 1307 bp in length
10792  10891:  gap of unknown length
10892  13010:  contig of 2119 bp in length
13011  13110:  gap of unknown length
13111  14617:  contig of 1507 bp in length
14618  14717:  gap of unknown length
14718  16217:  contig of 1500 bp in length
16218  16317:  gap of unknown length
16318  18106:  contig of 1789 bp in length
18107  18206:  gap of unknown length
18207  19498:  contig of 1292 bp in length
19499  19598:  gap of unknown length
19599  20867:  contig of 1269 bp in length
20868  20967:  gap of unknown length
20968  22203:  contig of 1236 bp in length
22204  22303:  gap of unknown length
22304  24098:  contig of 1795 bp in length
24099  24198:  gap of unknown length
24199  26005:  contig of 1807 bp in length
26006  26105:  gap of unknown length
26106  27441:  contig of 1336 bp in length
27442  27541:  gap of unknown length
27542  29037:  contig of 1496 bp in length
29038  29137:  gap of unknown length
29138  30660:  contig of 1523 bp in length
30661  30760:  gap of unknown length
30761  32755:  contig of 1995 bp in length
32756  32855:  gap of unknown length
32856  34483:  contig of 1628 bp in length
34484  34583:  gap of unknown length
34584  37578:  contig of 2995 bp in length
37579  37678:  gap of unknown length
37679  39699:  contig of 2021 bp in length
39700  39799:  gap of unknown length
39800  42411:  contig of 2612 bp in length
42412  42511:  gap of unknown length
42512  45638:  contig of 3127 bp in length
45639  45738:  gap of unknown length
45739  49739:  contig of 4001 bp in length
49740  49839:  gap of unknown length
49840  54276:  contig of 4437 bp in length
54277  54376:  gap of unknown length
54377  59051:  contig of 4675 bp in length
59052  59151:  gap of unknown length
59152  64626:  contig of 5475 bp in length
64627  64726:  gap of unknown length
64727  67988:  contig of 3262 bp in length
67989  68088:  gap of unknown length
68089  72494:  contig of 4406 bp in length
72495  72594:  gap of unknown length
72595  75643:  contig of 3049 bp in length
75644  75743:  gap of unknown length
75744  79510:  contig of 3767 bp in length
79511  79610:  gap of unknown length
79611  84664:  contig of 5054 bp in length
84665  84764:  gap of unknown length
84765  90514:  contig of 5750 bp in length
90515  90614:  gap of unknown length
90615  96185:  contig of 5571 bp in length
96186  96285:  gap of unknown length
96286 102048:  contig of 5763 bp in length
102049 102148: gap of unknown length
102149 108361: contig of 6213 bp in length
```

```
FEATURES             Location/Qualifiers
     * 108362 108461:  gap of unknown length
     * 108462 114329:  contig of 5868 bp in length
     * 114330 114429:  gap of unknown length
     * 114430 120644:  contig of 6215 bp in length
     * 120645 120744:  gap of unknown length
     * 120745 128025:  contig of 7281 bp in length
     * 128026 128125:  gap of unknown length
     * 128126 138788:  contig of 10663 bp in length
     * 138789 138888:  gap of unknown length
     * 138889 152044:  contig of 13156 bp in length.

     source          1..152044
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-45D20"
                     /clone_lib="RPCI-11"

BASE COUNT      40176 a   33928 c   33229 g   40362 t   4349 others
ORIGIN
```

```
RESULT 14
AC025947/c

LOCUS       AC025947   157057 bp   DNA      HTG      10-JUN-2000
DEFINITION  Homo sapiens chromosome 10 clone RP11-78A18, WORKING
            DRAFT SEQUENCE, 26 unordered pieces.
ACCESSION   AC025947
VERSION     AC025947.3  GI:8439851
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 157057)
  AUTHORS   Smith,D.R.
  TITLE     Genome Therapeutics Corporation Sequencing Center: Human Genome
            Sequence Data
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 157057)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02453, USA
COMMENT     On Jun 10, 2000 this sequence version replaced gi:7528340.
            -------------- Genome Center
            Center: Genome Therapeutics Corporation
            Center code: GTC
            Web site: http://www.genomecorp.com/
```

```
Query Match          3.5%;   Score 185.8;   DB 79;   Length 152044;
Best Local Similarity   80.8%;
Matches 232;  Conservative 0;  Mismatches 47;  Indels 8;  Gaps 1;

Qy   1321 acctggagctcactgcaacctctgccctcccaggttcaagcaattctccttgtccagcctcc 1380
          ||||| ||||||||||||||||| ||||||||||||||||| ||| ||||||||| |||||||
Db  30398 ATCTCGGCCTCACTGCAACCTCCACCTCCCGGGTTCAAGTGANTGATTCTCCTGCCTCAGCCTCC 30339

Qy   1381 cgcgtagctgggactacaggc---------gcacgcccggctaatttttgtattgttagta 1432
          | |||||||||||| |||||||         ||||||||| |||||||||||| ||| ||||
Db  30338 TGAGTAGCTGGGATTACAGGCGCACCCACTACCACGCCCAGCTAATTTTTGTAATTTTTAGTA 30279

Qy   1433 gagatggggtttcaccatgttagccgcggctggtcttgaacctcctgacctgatcca 1492
          |||||||||||||||||||||  || |||||||||||||||| |||||||| |||||
Db  30278 GAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCA 30219

Qy   1493 cccacctccagcctcctaaagtgctgggattacaggcgtgagccgcgcccggcaagg 1552
          |||| || |||||||| |||||||||||||||||||| |||||| | |||| |||||
Db  30218 CCCGCCTTAGCCTCCCAAAAGTGCTGGGATTACAGGCGATGAGCCACTGCACCTGGCCAATT 30159

Qy   1553 gtcagtgtttaataaggaataacttgaatggttactacaaccaacag 1599
          ||||||||||||||||||||||||||||| |||||||||||||||||
Db  30158 TTAATTTTTTTTTAATGAAAAAATGTGGATATATTTTAAAAAAAACACAG 30112
```

FEATURES

Contact: gtc-seqcenter@genomecorp.com
Center project name: hg307
Project Information
-----------Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Sequencing Program: Phrap; version 990315
Consensus quality: 140586 bases at least Q40
Consensus quality: 148596 bases at least Q30
Consensus quality: 150061 bases at least Q20
Insert size: 154557; sum-of-contigs
Quality coverage: 3.7x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
  be preserved.

* 1   1544: contig of 1544 bp in length
* 1545   1644: gap of unknown length
* 1645   3218: contig of 1574 bp in length
* 3219   3318: gap of unknown length
* 3319   5022: contig of 1704 bp in length
* 5023   5122: gap of unknown length
* 5123   6553: contig of 1431 bp in length
* 6554   6653: gap of unknown length
* 6654   8651: contig of 1998 bp in length
* 8652   8751: gap of unknown length
* 8752   10843: contig of 2092 bp in length
* 10844   10943: gap of unknown length
* 10944   13772: contig of 2829 bp in length
* 13773   13872: gap of unknown length
* 13873   15825: contig of 1953 bp in length
* 15826   15925: gap of unknown length
* 15926   19591: contig of 3666 bp in length
* 19592   19691: gap of unknown length
* 19692   22350: contig of 2659 bp in length
* 22351   22450: gap of unknown length
* 22451   25948: contig of 3498 bp in length
* 25949   26048: gap of unknown length
* 26049   29641: contig of 3593 bp in length
* 29642   29741: gap of unknown length
* 29742   34105: contig of 4364 bp in length
* 34106   34205: gap of unknown length
* 34206   38484: contig of 4279 bp in length
* 38485   38584: gap of unknown length
* 38585   41992: contig of 3408 bp in length
* 41993   42092: gap of unknown length
* 42093   47920: contig of 5828 bp in length
* 47921   48020: gap of unknown length
* 48021   54416: contig of 6396 bp in length
* 54417   54516: gap of unknown length
* 54517   60194: contig of 5678 bp in length
* 60195   60294: gap of unknown length
* 60295   67639: contig of 7345 bp in length
* 67640   67739: gap of unknown length
* 67740   74790: contig of 7051 bp in length
* 74791   74890: gap of unknown length
* 74891   85631: contig of 10741 bp in length
* 85632   85731: gap of unknown length
* 85732   93940: contig of 8209 bp in length
* 93941   94040: gap of unknown length
* 94041   104190: contig of 10151 bp in length
* 104191   104290: gap of unknown length
* 104291   116386: contig of 12096 bp in length
* 116387   116486: gap of unknown length
* 116487   135856: contig of 19370 bp in length
* 135857   135956: gap of unknown length
* 135957   157056: contig of 21100 bp in length.

Location/Qualifiers

source   1...157057
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="Chromosome 10"
                /clone_lib="RP11-78A18"
                /clone="RP11-11"

BASE COUNT   45178 a   32550 c   32372 g   44448 t   2509 others
ORIGIN

Query Match   3.5%;   Score 185.8; DB 78; Length 157057;
Best Local Similarity   80.8%;   Pred. No. 5.8e-30;
Matches 232; Conservative   0; Mismatches   47; Indels   8; Gaps   1;

Qy   1321 acctgagctcactgcaacctctggcctcccaaggtccaagcaattctccctgtctcagcctcc 1380
             ||||||||||||||||||||||||||||| ||||||||||||||||||||| ||||| |||||
Db   109551 ATCTCGGCCTCACTGCAACCTCCACCTCCACCTCGGGGTTCAAGTAGTATTCTCCTGCCTCAGCCTCC 109492

Qy   1381 cgcggtagctgggactacaggc----------gcacggccgggctaattttgttattgttagta 1432
             ||||||||||||||||| |||| |||||||||||||||||||||||||||||||
Db   109491 TGACTAGCTGGGATTACAGGCACCCACCACCACCACTACCGCCACTAATTTTGTATTTTTAGTA 109432

Qy   1433 gagatggggttcaccacacatattagcccggctggtcttgaactcctgacctcaggtgatcca 1492
             |||||||||||||||||||||||||||||||||||||| |||| ||||||||||| |||||
Db   109431 GAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCA 109372

Qy   1493 cccacctcagcctcccaaagtgctgggattacaggcatgagtcaccgcgcccggccaagg 1552
             ||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   109371 CCCGCCTTAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCACCTGGCCAATT 109312

Qy   1553 gtcagtgtttaataaggaataactttgaatggtgtttactaaacaaacag 1599
             |||||||||||||||||||||||||| ||||| |||| |||| |||
Db   109311 TTAATTTTTTTAATGAAAAATGTGGATATTTTAAAAAAAACACAG 109265

RESULT 15
AC015488
LOCUS   AC015488   161499 bp   DNA   HTG   16-MAR-2000
DEFINITION   Homo sapiens clone RP11-20F6, WORKING DRAFT SEQUENCE, 16 unordered
                pieces.
ACCESSION   AC015488
VERSION   AC015488.4 GI:7107977
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   human
ORGANISM   Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 161499)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE   Homo sapiens, clone RP11-20F6
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 161499)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.,
                Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
                Brown,A., Castle,A., Colangelo,N., Collins,S., Collymore,A.,
                Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
                Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
                Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
                Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
                Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
                McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
                Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
                Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
                Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
                Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
                Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE   Direct Submission
JOURNAL   Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
                Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Feb 28, 2000 this sequence version replaced gi:671593.
                All repeats were identified using RepeatMasker:
                Smit, A.F.A. & Green, P. (1996-1997)
                http://ftp.genome.washington.edu/RM/RepeatMasker.html

```
-------- Genome Center
        Center: Whitehead Institute/ MIT Center for Genome Research
        Center code: WIBR
        Web site: http://www-seq.wi.mit.edu
        Contact: sequence_submissions@genome.wi.mit.edu
-------- Project Information
        Center project name: L3895
        Center clone name: 20_F_6
-------- Summary Statistics
        Sequencing Vector: M13; M77815; 100% of reads
        Chemistry: Dye-terminator Big Dye; 100% of reads
        Assembly program: Phrap; version 0.960731
        Consensus quality: 152793 bases at least Q40
        Consensus quality: 156653 bases at least Q30
        Consensus quality: 158136 bases at least Q20
        Insert project name: L3895
        Insert size: 160000; agarose-fp
        Insert size: 159999; sum-of-contigs
        Quality coverage: 5.9 in Q20 bases; agarose-fp
        Quality coverage: 5.9 in Q20 bases; sum-of-contigs
--------
        NOTE:This is a 'working draft' sequence. It currently
        consists of 16 contigs. The true order of the pieces
        is not known and their order in this sequence record is
        arbitrary. Gaps between the contigs are represented as
        runs of N, but the exact sizes of the gaps are unknown.
        This record will be updated with the finished sequence
        as soon as it is available and the accession number will
        be preserved.
--------
*        1   2206: contig of 2206 bp in length
*     2207   2306: gap of 100 bp
*     2307   6910: contig of 4604 bp in length
*     6911   7010: gap of 100 bp
*     7011   9818: contig of 2808 bp in length
*     9819   9918: gap of 100 bp
*     9919  13394: contig of 3476 bp in length
*    13395  13494: gap of 100 bp
*    13495  17122: contig of 3628 bp in length
*    17123  17222: gap of 100 bp
*    17223  19752: contig of 2530 bp in length
*    19753  19852: gap of 100 bp
*    19853  23958: contig of 4106 bp in length
*    23959  24058: gap of 100 bp
*    24059  29684: contig of 5626 bp in length
*    29685  29784: gap of 100 bp
*    29785  33951: contig of 4167 bp in length
*    33952  34051: gap of 100 bp
*    34052  39841: contig of 5790 bp in length
*    39842  39941: gap of 100 bp
*    39942  48297: contig of 8356 bp in length
*    48298  48397: gap of 100 bp
*    48398  56898: contig of 8501 bp in length
*    56899  56998: gap of 100 bp
*    56999  65924: contig of 8926 bp in length
*    65925  66024: gap of 100 bp
*    66025  81114: contig of 15090 bp in length
*    81115  81214: gap of 100 bp
*    81215 112500: contig of 31286 bp in length
*   112501 112600: gap of 100 bp
*   112601 161499: contig of 48899 bp in length.

FEATURES            Location/Qualifiers
    source          1..161499
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="RP11-20F6"
                    /clone_lib="RPCI-11 Human Male BAC"
    misc_feature    1..206
                    /note="assembly_fragment"
    misc_feature    2307..6910
                    /note="assembly_fragment"
    misc_feature    7011..9818
                    /note="assembly_fragment"
    misc_feature    9919..13394
                    /note="assembly_fragment"
    misc_feature    13495..17122
                    /note="assembly_fragment"
    misc_feature    17223..19752
                    /note="assembly_fragment"
    misc_feature    19853..23958
                    /note="assembly_fragment"
    misc_feature    24059..29684
                    /note="assembly_fragment"
    misc_feature    29785..33951
                    /note="assembly_fragment"
    misc_feature    34052..39841
                    /note="assembly_fragment"
    misc_feature    39942..48297
                    /note="assembly_fragment"
    misc_feature    48398..56898
                    /note="assembly_fragment"
    misc_feature    56999..65924
                    /note="assembly_fragment"
    misc_feature    66025..81114
                    /note="assembly_fragment"
                    /clone_end:SP6
                    /vector_side:left"
    misc_feature    81215..112500
                    /note="assembly_fragment"
    misc_feature    112601..161499
                    /note="assembly_fragment"
                    /clone_end:T7
                    vector_side:right"
    misc_feature    13495..17122
                    /note="assembly_fragment"
    misc_feature    17223..19752
                    /note="assembly_fragment"

BASE COUNT    44611 a    37543 c    36473 g    41370 t    1502 others
ORIGIN

Query Match              3.5%;     Score 185.8;     DB 72;     Length 161499;
Best Local Similarity   77.1%;     Pred. No. 5.8e-30;
Matches 226;     Conservative   0;     Mismatches  67;     Indels   0;     Gaps   0;

Qy   1283 gtgagggtctgtgtcttacacctacctgtatgctctacacctgagctcactgcaacctct 1342
          |||||||| ||||||||||||||||||||||||||||| ||| ||||||||||||||
Db  16769 GTCTTGCTCTGTCACCCAGGCTGGAGGGCAGTGGCACCATCTCGGCTCACTGCAACCTCT 16828

Qy   1343 gcctcccaggttcaagcaattctcctgtcctcagcctcccgcgtagctgggactacaggcg 1402
          ||||||||||| |||||||||||||||||||||||| |||| ||||||||||||||
Db  16829 GCCTCCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTACCTGGGGTTATAGGCA 16888

Qy   1403 cacgcccgcgctaatttttgtattgttagtagaagaatggggtttcaccatattagccggct 1462
          ||||| || ||||||||||||| || ||||||||| | |||||||||||||||
Db  16889 CATGCCAGGCTAATTTTTGTGTTTTTAGTAGAGATGGGGTTTGCCATGTTGGCCAGGCT 16948

Qy   1463 ggtcttgaactcctgacctccaggtgatccaccccaacctcagcctcctaaagtctgggatt 1522
          ||||||||||||||||||||||||||||||| |||| |||||||||||||||||
Db  16949 GGTCTTGAACTCCTGACCTCAGATGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATT 17008

Qy   1523 acaggcatgagtcaccgcgcccggcaagggtcagtgttaataagggaataac 1575
          ||||| ||| |||||||||||||||||||||| | | ||||| |||
Db  17009 ACAGGTGTGAGCCACCGCGCCCAGCCTATTTTTAATTTTTATTTTTATCAAC 17061

Search completed: December  3, 2000, 19:17:56
Job time: 46490 sec
```